MAR 2 6 1928

VOLUME XXIII NEW SERIES, NO. 161

JOURNAL
OF THE
AMERICAN STATISTICAL ASSOCIATION

MARCH, 1928

CONTENTS

THE RÔLE OF STATISTICS IN BUSINESS FORECASTING. By Edmund E. Day 1

INDEX NUMBER BIAS. By W. V. Lovitt 10

THE POPULATION OF CHINA IN 1910. By Walter F. Willcox 18

THE STANDARD DEVIATION OF THE CORRELATION COEFFICIENT. By A. L. Bowley 31

AN EXPRESSION FOR THE ASSYMMETRICAL TENDENCY OF FREQUENCY DISTRIBUTIONS. By Stephen Varga 35

A STATISTICAL STUDY OF THE ECONOMIC AND SOCIOLOGICAL SIGNIFICANCE OF POPULATION AGE. By Nathan P. Levin ... 40

AN AGGREGATIVE INDEX OF FARM PURCHASING POWER. By Forrest Bee Ashby 49

NOTES:

An Elementary Proof of Stirling's Formula. By Robert E. Moritz 55

Apparatus to Facilitate the Calculation of the Moments of a Distribution. By Thomas N. Jenkins 58

Average Variations in Yields, Acreage and Production. By L. J. Norton 61

Occupations of the Mexican-Born Population of Texas, New Mexico, and Arizona, 1900–1920. By Roden Fuller 64

Price Trends and Prosperity (68); New Methods of Measuring Public Health Progress (72); Progress of the Census (76); Some Suggestions for Securing Standardized Criminal Statistics (77); A Standard Reporting System for Public Accidents (79); Miscellaneous Notes (81); Members Added (86).

REVIEWS 89

RECENT LITERATURE 110

PUBLISHED QUARTERLY BY THE
AMERICAN STATISTICAL ASSOCIATION
PUBLICATION OFFICE: Rumford Press, Concord, N. H.
EDITORIAL OFFICE: Columbia University, New York City

Price $1.50 per copy **$5.00 per annum**

THE EMPLOYMENT CLEARING HOUSE
OF THE
AMERICAN STATISTICAL ASSOCIATION

will aid you with your employment problem. It makes no charge to employers for its services. Any member of the Association may register for a position. There is no fee.

The following positions were open since December, 1927:

Field	Pay									
	Under $1800		$1800–2400		$2500–3600		$3700–4900		$5000 and over	
	M	F	M	F	M	F	M	F	M	F
Agricultural Economics										
Business Forecasting			1							
Business—Mfg. and Selling		1	1		1					
Drafting				1						
Economic and Industrial Research	2	4			2		1		1	
Governmental	1	1								
Merchandising					1					
Social Work				1						
Statistical Secretarial										
Technical Computing										
Teaching							1			

Members were looking for openings as listed:

Field	Pay									
	Under $1800		$1800–2400		$2500–3600		$3700–4900		$5000–and over	
	M	F	M	F	M	F	M	F	M	F
Agricultural Economics					1					
Business—Mfg. and Selling	1		1	2	10		1		5	
Economic and Indus. Research				2			1		1	
Finance	2	3	3		3	3	5		3	
Insurance	2	1				1	2			
Marketing					5	2	2			
Social Research				2	1	1				1
Teaching	1			2			1			
Transportation					4				1	
Vital Statistics					1					

COMMUNICATE YOUR NEEDS TO

HELEN W. LINKE

Secretary's Assistant, American Statistical Association

Room 530, Commerce Bldg., 236 Wooster St., New York, N. Y.

[NOT A PAID ADVERTISEMENT]

NEW SERIES, NO. 161 (VOL. XXIII) MARCH, 1928

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Formerly the Quarterly Publication of the American Statistical Association

CONTENTS

THE RÔLE OF STATISTICS IN BUSINESS FORECASTING. By EDMUND E. DAY 1

INDEX NUMBER BIAS. By W. V. LOVITT 10

THE POPULATION OF CHINA IN 1910. By WALTER F. WILLCOX 18

THE STANDARD DEVIATION OF THE CORRELATION COEFFICIENT. By A. L. BOWLEY 31

AN EXPRESSION FOR THE ASYMMETRICAL TENDENCY OF FREQUENCY DISTRIBUTIONS. By STEPHEN VARGA 35

A STATISTICAL STUDY OF THE ECONOMIC AND SOCIOLOGICAL SIGNIFICANCE OF POPULATION AGE. By NATHAN P. LEVIN 40

AN AGGREGATIVE INDEX OF FARM PURCHASING POWER. By FORREST BEE ASHBY 49

NOTES:

An Elementary Proof of Stirling's Formula. By Robert E. Moritz 55

Apparatus to Facilitate the Calculation of the Moments of a Distribution. By Thomas N. Jenkins 58

Average Variations in Yields, Acreage and Production. By L. J. Norton 61

Occupations of the Mexican-Born Population of Texas, New Mexico, and Arizona, 1900–1920. By Roden Fuller 64

Price Trends and Prosperity (68); New Methods of Measuring Public Health Progress (72); Progress of the Census (76); Some Suggestions for Securing Standardized Criminal Statistics (77); A Standard Reporting System for Public Accidents (79); Miscellaneous Notes (81); Members Added (86).

REVIEWS: Mitchell: *Business Cycles: The Problem and Its Setting*, by H. Working (89); Vanderblue and Crum: *The Iron Industry in Prosperity and Depression*, by W. L. Thorp (94); National Industrial Conference Board: *Workmen's Compensation Problem in New York State*, by R. H. Blanchard (97); Artman: *Food Costs and City Consumers*, by G. E. Barnett (100); Dahlberg: *Twin Births and Twins from a Hereditary Point of View*, by C. B. Davenport (101); Collins: *Economic Status and Health*, by I. M. Rubinow (102); Angell: *The Theory of International Prices: History, Criticism and Restatement*, by J. H. Hollander (104); Hall: *Banking Cycles*, by R. W. Burgess and W. R. Burgess (106); Bodfish: *Money Lending Practices of Building and Loan Associations in Ohio*, by J. G. Rolph (107); Lagerquist: *Public Utility Finance*, by H. K. Herwitz (108); Britten and Thompson: *A Health Study of Ten Thousand Male Industrial Workers*, by R. Henderson (109).

RECENT LITERATURE 110

Published Quarterly by the American Statistical Association

RUMFORD PRESS BUILDING, CONCORD, N. H.

Editorial Office, Columbia University, New York City

Entered at the post-office, Concord, N. H., as second-class mail matter

Acceptance for mailing at the special rate of postage provided for in Section No. 1103, Act of October 3, 1917, authorized April 29, 1922

Officers of the American Statistical Association

ORGANIZED NOVEMBER 27, 1839

President, Carl Snyder

Vice-Presidents, Seymour L. Andrew, Raymond Pearl, Ralph G. Hurlin

Counsellors, James W. Glover, W. Randolph Burgess, R. H. Coats

Secretary-Treasurer, Willford I. King,
Room 530, Commerce Bldg., New York University,
236 Wooster Street, New York, N. Y.

District Secretaries, Robert W. Bachelor,
1437 Bancroft Way, Berkeley, Calif.

F. L. Carmichael,
University of Denver, Denver, Colo.

George J. Eberle,
810 South Spring Street, Los Angeles, Calif.

H. B. Flinkers,
Federal Reserve Bank of Cleveland, Ohio

Dorsey W. Hyde, Jr., Secretary,
Chamber of Commerce of Washington, D. C.

John H. Cover,
University of Pittsburgh, Pittsburgh, Pa.

A. Heath Onthank,
George Harrison Phelps, Inc., Detroit, Mich.

Roswell F. Phelps,
Division of Statistics, Massachusetts Dept. of Labor and Industries, Boston, Mass.

S. W. Russell,
Commercial Research Department,
Swift and Company, Chicago, Ill.

C. D. Simmons,
University of Texas, Austin, Texas

Harry S. Will,
66 Fallis Road, Columbus, Ohio

Editor, Frank Alexander Ross,
Columbia University, New York City

Associate Editors, Edmund E. Day, Wesley C. Mitchell, William F. Ogburn, Walter F. Willcox

Leo Wolman,
30 Fifth Avenue, New York City

NEW SERIES, NO. 161 (VOL. XXIII) MARCH, 1928

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Formerly the Quarterly Publication of the American Statistical Association

THE RÔLE OF STATISTICS IN BUSINESS FORECASTING [1]

By Edmund E. Day

In any large group of social scientists—such as the one meeting here this evening—widely divergent attitudes toward business forecasting are certain to be represented. At one extreme will be the view that business forecasting, though doubtless an unavoidable phase of the conduct of practical affairs, has as yet no scientific standing. According to this view, business forecasting when given scientific form is to be regarded as a sort of malpractice, not without its adherents and for the time being distinctly lucrative, but certainly not entitled to recognition by legitimate scientific practitioners. At the other extreme will be the opinion that in business forecasting is found one of the most valuable lines of endeavor in the entire field of social science—a line in which realism, objectivity, and adequacy of treatment are compelled by the very nature of the problems encountered; a line in which weaknesses of procedure are most mercilessly exposed and high standards of performance most effectively enforced. In between these extreme views will lie attitudes ranging all the way from mild skepticism on the one hand to benign and hopeful interest on the other. Clearly there is no consensus of opinion as to the place business forecasting should be assigned in the evolving program of social science.

Into this diversity of opinion at least one proposition may be profitably projected: Statistics and business forecasting have effected connections which are not likely to be speedily broken. It is unnecessary to determine here which one of the two made the first approaches; nor is it worth while to scrutinize the motives which led to the union. Doubtless there are those who will regard the affair as a "marriage of convenience," dictated by obvious pecuniary advantages; but that need not concern us at the present time. The important fact

[1] Presidential address at the Eighty-Ninth Annual Meeting of the American Statistical Association at Washington, D. C., December 28, 1927.

is that the attachment of business forecasting and statistics rests upon no passing fancy. The two are assuredly joined for life. The statistical clan, consequently, cannot afford to regard the situation with indifference. The family reputation is too clearly at stake. A conclave to give the entire matter thoughtful and sympathetic consideration would seem to be in order.

The relation of statistics to business forecasting can best be analyzed if at the outset there is definite recognition of three general types of prediction: (1) prediction by analogy; (2) prediction by formula; (3) prediction by analysis. These three types rarely, if ever, appear in pure or unmixed examples. Yet they represent such distinct differences in the logic of business forecasting that it is only in terms of each type in turn that the relation of statistics to forecasting is likely to be clarified.

In the first of the three forms, prediction is effected through resort to the argument from analogy. It is reasoned that the present period is like a certain period in the past and that, therefore, the events lying just ahead will be like those which followed in the past period. Of course, recognition of these simila.ities and dissimilarities of experience lies at the very foundation of all scientific inquiry. But in prediction by analogy the resemblances which are made the basis of prediction are of a general order. Comparison does not take the form of detailed analysis. The forecast deals with the situation in the large.

An illustration of this type of prediction is to be found in certain forecasts of the present post-war price trend. Ready reference has been made in these predictions to monetary experiences following certain earlier wars—the Napoleonic War in Europe and the Civil War in the United States are the two most commonly cited. It is argued that, inflation having attained certain heights and receded during certain periods in connection with these earlier occasions of armed conflict on a large scale, inflation may be expected to run about the same course in connection with the recent World War. In other words, the earlier post-war experiences are lifted over bodily. Time curves representing the price movement of the earlier period may be superimposed graphically on the later to give the prediction more concrete form. Other examples of this type of forecasting could be readily supplied, but the type is too familiar to require further illustration.

In forecasting of this sort the statistics that are employed merely serve to give the prediction more definite form; the underlying basis of the forecasting—namely, the assumption that the two periods in question, insofar as the elements related to the point at issue are concerned, are essentially alike—is not developed statistically. Predic-

tion by analogy makes connection with statistics, but its relation to statistics is superficial, not essential.

Prediction by analogy ordinarily falls of its own manifest superficiality. To observe that monetary inflation in connection with the Civil War came to an end about thirteen years after the close of hostilities and to conclude from this that pre-war prices will be restored in the United States in our own time, say in 1931, is to deal so exclusively with surface resemblances and so little with underlying differences as to discredit any conclusion before it is presented. In social investigation, argument from analogy has its place. At times it is highly suggestive. But it is never conclusive. Business forecasting which proceeds entirely by analogy is not entitled to serious consideration.

Prediction by formula is a more interesting and significant type. The primary characteristic of the type is to be found in the discovery of specific critical factors, the condition and previous course of which determine events that are to ensue. A formula is developed to describe the behavior of the variable, or variables, in question. With this formula in hand, prediction becomes purely a mathematical process.

This general type of forecasting may be most profitably examined in certain concrete varieties. One of the most interesting of these consists of the prediction of the future values of numerous aggregates—such as city population, automobile registration, cigarette consumption—through the application of a normal growth curve. Here the future course of events is supposed to be forecast by the rate of change exhibited by the variable in its manifestations up to date. A curve having been fitted to the data on some formula which appears both to yield a reasonably close fit and to conform to certain notions of the underlying situation, prediction takes the form of simple extrapolation. A general law of growth, or of increase, is postulated. The procedure throughout is highly statistical. The prediction derives directly from a mathematical function.

There can be no question regarding the value of this form of analysis in certain cases. When the variable of which future values are sought consists of a great aggregate of individual items the vast majority of which are practically certain to carry over, there is in the mass as a whole a stability or inertia which greatly facilitates the forecasting process. Such statistical masses are not often subject to sudden radical changes of size. They move forward in deep channels which render highly unlikely any sharp diversions. Variables of this character offer the most favorable opportunities for prediction by formula.

Certain grounds for caution nevertheless are to be noted. In the first place there is not yet any convincing case for the exclusive use of

a single formula for the so-called "law of growth." True, any formula that is to be considered favorably must give expression to varying rates of increase during certain distinct stages in the characteristic development of observed mass phenomena. But a whole family of curves will meet this requirement. Selection from among this family can hardly be made on the basis of "closeness of fit," since inaccuracies of the empirical data are ordinarily of a higher order than are the failures of fit. At the present time one must conclude that there is no single formula for the "law of growth." In other words there is still room for differences of competent opinion as to just what is meant by the "law of growth."

In the second place it must be recognized that prediction through the application of a curve of growth is attended by obvious uncertainty when the variable in question lacks inherent inertia. Thus, present predictions of the demand for automobiles in the United States, based upon a growth curve fitted to the past demand for cars, may be quite upset by the general adoption of multiple-car ownership. Statistical masses of great magnitude are sometimes radically affected by new forces of an unforeseen and essentially unpredictable sort; statistical variables not presented in the form of continuing masses are obviously even more susceptible to sudden shifts: Witness the many disruptions of trade occasioned by the Great War, and the more recent unhappy sales experience of certain of the textile industries under the influence of changing styles.

Finally, it is to be noted that not infrequently in the fitting of curves of growth, or trend lines, future values are implied which challenge belief in the validity of the analysis for any considerable period ahead. In general, prediction through the use of a curve of growth is warranted to just the extent to which a thorough analysis of the factors underlying growth establish the reasonableness of the curve. For every mathematical formulation an adequate rational basis must be supplied. But in meeting this requirement, prediction by formula is almost certain to become prediction through analysis—quite a different matter.

Another variety of prediction by formula is to be found in those forecasts which rest upon the notion of periodicity in the business cycle; or, upon what amounts practically to the same thing, the notion of constant duration in the several phases of the cycle. Given the periodogram in the one case, or the standard period in the other, the problem of forecasting becomes a matter of locating certain basing points and of applying thereto the appropriate quantitative measure. It is only fair to say that there are few, if any, reputable statisticians who would recognize their own procedure in any such bald description. Yet not a few busi-

ness forecasts are the product of reasoning in which the idea of relatively fixed periodicity in one guise or another plays the leading part. Increasing knowledge of the business cycle, however, is making the concept of fixed periodicity appear more and more unreal. Duration of phases is itself one of the variables with which cycle analysis has to deal. No periodic function, however carefully derived, will solve the problem of business forecasting.

Still another variety of prediction by formula may be cited: Namely the prediction that rests upon specified critical values, or turning points, of some particular element or ratio extracted from the general business situation. Numerous examples of this sort of forecast could be readily cited. In appraising forecasts of this kind it is to be conceded that, not infrequently, the variable which is made the basis of the forecast is highly significant. The weakness of the prediction lies not in the irrelevance of the citation but in its inadequacy. No single factor, however important, can be made the basis of scientific business forecasting.

The ear-mark of all prediction by formula is complete reliance upon previously devised quantitative measures of factors which are conceived to be definitely determinative of events to come. The underlying concepts are essentially mechanistic; the formulation is characteristically mathematical. Laws are commonly postulated. Specific elements are emphasized to the exclusion of all else. Great reliance is frequently placed upon high coefficients of correlation without recognition of the fact that a coefficient of correlation may be so high as almost surely to indicate nominal relationship of the variables or manipulative treatment of the data. As might be expected, the results of prediction by formula have been unfortunate. Few things have done more to discredit statistics than the use of this sort of statistical procedure in business forecasting; few things more dangerous in forecasting than a statistician at large with a pet formula.

Prediction by analysis differs from prediction by formula not so much in its outward form as in its general philosophy. It makes extensive use of quantitative methods but does not exclude from consideration elements which at least for the present cannot be given quantitative expression. It throws into precise mathematical form as many of its analyses as lend themselves to such treatment, but it does not regard the resultant mathematical formulations as anything more than serviceable symbolic expressions of the underlying concrete relationships. In the eyes of prediction by analysis the significance of a formula is to be found only in the presumptive reality and reasonableness of the relationship the formula sets forth. Furthermore, for prediction by

analysis, solution of the problem of business forecasting is to be found in *continuous* study of total business situations with reference particularly to the discovery of significant connections between variables. Finally, prediction by analysis proceeds always with a minimum of preconception and a maximum of scientific open-mindedness. Here are elements which cut clear through the relationship of statistics to business forecasting.

A thorough understanding of what is meant by prediction by analysis presupposes a more general understanding of the nature of statistical methods and of statistical results. In this connection certain fundamental facts need to be clearly recognized. In the first place, a distinction must be made between description or perception, and inference. Statistical method is *par excellence* a means of rendering description more accurate and illuminating. To the extent that the subject matter is complex—in other words, to the extent that a multiplicity of forces join to bring about a given result—statistical methods are indispensable. The rapidly increasing knowledge of the business cycle which has been obtained in this country in recent years is the direct result of statistical inquiry. Original data, registering movements of a composite sort, have been resolved into their simpler elements. Trends, seasonal variations, cyclical fluctuations, irregular movements have all been brought clearly to light. The work has been highly significant. Nevertheless it has been essentially descriptive in character.

In recognizing the descriptive character of many statistical results, one should not conclude that the results derive entirely from the original observations. The statistician plays an important part even in descriptive statistics. For example, consider the task of constructing an index of the course of wholesale commodity prices in the United States since 1914. Here the final figures are designed to be merely an accurate record of past events—a descriptive account of important economic changes. But what the final figures are depends in very considerable measure upon what the statistician does—upon the data he selects, the averaging methods he pursues, the weighting principles he adopts. There is no more common fallacy regarding statistics than this: That statistics can be safely divorced from the statistician. In statistical work, even of a descriptive sort, the acceptability of the results depends upon the credibility of the worker.

A like statement is even more to the point when statistical work ceases to be primarily descriptive and becomes essentially analytical and rational. Here we deal with the processes of inference. When correlation coefficients are computed, the statistician rarely confines himself—as indeed he should not—to a mere quantitative statement of

the results. Inferences are drawn as to the relationship of the variables. These inferences commonly represent a bold adventure into regions by no means commanded by the statistical artillery which has been brought to bear. Frequently the statistical results are wholly consistent with any number of widely divergent inferences. True, the quantitative evidences, if carefully wrought, may make certain conclusions highly improbable if not altogether untenable. But statistical results have their limitations. To quote from Mr. R. A. Fisher: "If we choose a group of social phenomena with no antecedent knowledge of the causation or absence of causation among them, then the calculation of correlation coefficients, total or partial, will not advance us a step toward evaluating the importance of the causes at work." Inferences which by no means necessarily derive from the data form an integral part of most intricate statistical work. As a result, statistical analyses commonly expose the character of the statistician as definitely as they do that of the original statistical data.

It should not be concluded from these statements that the contributions of statistics are to be regarded as of doubtful value. After all, there are competent, cautious, reputable statisticians. In the hands of these, statistical methods yield highly significant results. Accurate description of phenomena, especially if they be phenomena presented in an intricate and complicated record of both relevant and irrelevant variables, constitutes an indispensable phase of scientific inquiry. Great strides are to be made in all lines of social investigation through improved and extended analyses in which description and inference join forces under intelligent supervision. Statistical analysis, even if its results be accepted always with reservations, has a most important contribution to make to social science.

A further fundamental feature of statistical methods is to be noted. Statistical results take form in a universe in which variation, approximation, contingency, and probability are prevailing conditions. Laboratory methods as applied in the physical sciences may reduce uncontrolled variations to negligible proportions; but in social situations, the isolation of factors is quite impossible. The partially known and the unknown unite to fix the record of social phenomena. As a rule, the social scientist is confronted with situations into which material, economic, political, social, and psychological forces play with extraordinary complexity. Were it not for the influences of custom, established order, habit, suggestion, inertia, and the like, the task of prediction would present insurmountable difficulties. As it is, forecasts in the field of social science are always to be accepted subject to certain limitations.

These limitations are imposed by a number of factors. First, many of the forecasts rest upon averages behind which wide variations in the individual items may be concealed. This fact of individual variation is fundamental. Second, most forecasts involve the use of data which are mere approximations. The element of error is omnipresent. Third, prediction in all social science is governed by contingency, not certainty, and the conclusions can only be stated as more or less probably correct. There will always remain a measure of art in business forecasting. Breadth of general information and solidity of business judgment will always make their intangible but none the less important contributions. These factors set serious limitations upon business forecasting; they do not destroy its usefulness nor its scientific standing.

With these ideas regarding the general nature of statistical methods and of statistical results clearly in mind, we may return to the consideration of prediction by analysis. That this type of prediction will make extensive use of statistics is obvious enough; only by resort to quantitative methods is prediction to be made at all scientific. Statistical analysis focussed on the time relationship of variables is certain to be especially valuable, exposing what may appear to be the "prime mover" in a given situation, or disclosing factors which may behave in such a fashion as to suggest joint causation or mutual interaction. Studies of association and correlation in time are bound to constitute the backbone of statistical method as applied to business forecasting. Along this line there is a task of major proportion, and one promising substantial results.

But statistical analysis even of this more constructive sort is not all that is necessary to scientific business forecasting. Thoroughgoing theoretical analysis—economic, political, sociological, social—realistically conceived and constructively applied, must be joined with every important step in the statistical procedure. For too long the statistician and the theorist have gone their individual ways. They should become inseparable companions. If together they face the task of business forecasting as one worthy of their talents, as a task requiring continuous application and never-failing caution, the relations of business forecasting and statistics need cause no concern. Increasingly valuable results may be confidently awaited.

Meanwhile it may not be amiss to mention briefly in closing what appear to be the chief conditioning factors on the statistical side in the further progress of the work.

In the first place, it is plainly evident that discoveries of new statistical technique are not just now a pressing need. Not that improvements of technique are to be neglected; they are desirable at all times, of course. But the improvement of business forecasting does not now

await the development of more refined and precise instruments of statistical analysis. Adequate apparatus is already at hand. Progress depends upon its more effective employment.

The importance of more comprehensive analysis, on the other hand, can hardly be exaggerated. We need to know more about consumers' purchasing habits, family incomes, the accumulation of surplus funds, margins of profit, retail and wholesale stocks, price-determining factors, interactions of credit, to mention only a few of the factors about which we now know all too little. In many instances the need is for additional original data. In other cases, existing data require more thorough analysis. A large number of relationships have yet to be subjected to statistical measurement. In general, business forecasting should be given a much broader and deeper statistical foundation than has thus far been provided.

Finally, there is every reason for accompanying statistical work in the field of business forecasting with a continuous appraisal of results, made generally available through publication of the findings. The failures are not necessarily discreditable; in considerable part they are for the present unavoidable. Why should they not be frankly acknowledged, and as fully as possible dispassionately explained? Out of these mistakes of the past, valuable lessons for the future are to be learned. The difficulties in any course of free confession are to be acknowledged, of course. I can almost hear the words: What of the subscribers? What of the board? What of this vice-president, or that? But if business forecasting is to achieve scientific status it must free itself of the restraints imposed by unenlightened commercial interests. This may sound like an impracticable counsel of perfection. Nevertheless it is offered seriously. Business forecasting should profit from the fact that its results are verifiable; it should be given constant evaluation in the light of subsequent events. Only so is the work likely to go forward as it should with the greatest possible accumulation of knowledge and insight.

Reverting to the differences of attitude with which these remarks were introduced, one final conclusion may be ventured. The connections of business forecasting and statistics can no longer be fairly regarded either with intolerance or with indifference. There are in the possible contributions of statistical methods to business forecasting a significance and a promise that challenge the utmost that social science has to offer. Let us as scientists accept this challenge; accept it seriously and with courage, but with a humility born of the manifest limitations of our methods and of the extraordinary complexity of the situations with which we have to deal.

INDEX NUMBER BIAS

By W. V. Lovitt, *Colorado College*

One requisite of a good index number is that it shall satisfy the time reversal test, which requires that the product of the index number for any given year on a base year by the index number for the base year on the given year shall be unity. Thus, if an article sells in Havana for forty cents and in New York for sixty cents the index number of price with Havana for base is 3/2; the index number with New York for base is 2/3; the product of 3/2 and 2/3 is unity. For one article, with prices unweighted, this test is always satisfied.

If this product is greater than unity, the index number is said to have an upward bias. If this product is less than unity, the index number is said to have a downward bias.

It is well known that the geometric average of relatives satisfies the time reversal test and thus has no bias. The arithmetic and harmonic averages of relatives do not satisfy the time reversal test and thus have what is called *type* bias.

Let us use the usual notation,

p^i_0 = price of the i^{-th} commodity in the base year (say 1913)
p^i_1 = price of the i^{-th} commodity in the given year (say 1927)
q^i_0 = quantity of the i^{-th} commodity in the base year
q^i_1 = quantity of the i^{-th} commodity in the given year

For the unweighted arithmetic average of relatives, if this average has an upward bias, we have to show that

$$1 < \frac{\sum_{i=1}^{N} \frac{p^i_1}{p^i_0}}{N} \cdot \frac{\sum_{j=1}^{N} \frac{p^j_0}{p^j_1}}{N}.$$

This inequality readily reduces to

$$0 < \sum_{i=1}^{N} \sum_{j=1}^{N} (p^i_0 p^j_1 - p^j_0 p^i_1)^2 \prod_{s=1}^{N} p^s_0 \cdot \prod_{s=1}^{N} p^s_1 \quad (i \neq j;\ s \neq i, j).$$

Since the square of any real number is positive and all prices are positive, every term on the right is a positive term, their sum is positive and therefore greater than zero. This establishes the upward bias of the unweighted arithmetic average of relatives.

For the unweighted harmonic average of relatives, we have, if this average has a downward bias, to show that

$$\frac{N}{\Sigma\frac{p_0}{p_1}} \cdot \frac{N}{\Sigma\frac{p_1}{p_0}} < 1$$

or

$$N^2 < \Sigma\frac{p_0}{p_1} \cdot \Sigma\frac{p_1}{p_0}.$$

This is the same inequality which we met in discussing above the unweighted arithmetic average of relatives. Thus we establish the downward bias of the unweighted harmonic average of relatives.

The four following methods of weighting by values are in current use:

I. Each weight equals base year price times base year quantity (p_0q_0)
II. Each weight equals base year price times given year quantity (p_0q_1)
III. Each weight equals base year price times base year quantity (p_1q_0)
IV. Each weight equals given year price times given year quantity (p_1q_1)

Let us apply these quantities as weights in a weighted arithmetic average of relatives.

Using I and applying the time reversal test we have to discover whether

$$A \equiv \frac{\Sigma p_1q_0}{\Sigma p_0q_0} \cdot \frac{\Sigma p_0q_1}{\Sigma p_1q_1} \gtreqless 1.$$

This inequality readily reduces to

$$0 \gtreqless \Sigma\Sigma(p^i{}_0p^j{}_1 - p^j{}_0p^i{}_1)(q^i{}_0q^j{}_1 - q^j{}_0q^i{}_1). \tag{1}$$

Whence it would seem that there is no necessary fixed relationship. That this is true is readily shown by numerical illustrations. Thus for the set of values

	p_0	q_0	p_1	q_1
(a)	1	2	3	1
	1	3	2	2

we find

$$A = \frac{36}{35} > 1.$$

For the set of values

	p_0	q_0	p_1	q_1
(β)	1	2	3	1
	1	3	2	1

we find

$$A=\frac{24}{25}<1.$$

For the set of values

	p_0	q_0	p_1	q_1
(γ)	1	2	3	1
	1	3	2	1.5

we find

$$A=1.$$

Thus any statement that this index number has any fixed upward or downward bias is erroneous.

Using II and applying the time reversal test we have to discover whether

$$B\equiv\frac{\Sigma p_0q_1\dfrac{p_1}{p_0}}{\Sigma p_0q_1}\cdot\frac{\Sigma p_1q_0\dfrac{p_0}{p_1}}{\Sigma p_1q_0}=\frac{\Sigma p_1q_1}{\Sigma p_0q_1}\cdot\frac{\Sigma p_0q_0}{\Sigma p_1q_0}\lessgtr 1.$$

This inequality readily reduces to (1) above. Whence it is seen that no necessary fixed relationship exists. We illustrate this by the following sets of values.

Thus for (β) we find $B=\frac{25}{24}>1.$

For the set of values (α) we find $B=\frac{35}{36}<1.$

For the set of values (γ) we find $B=1.$

Using III and applying the time reversal test we have to discover whether

$$C\equiv\frac{\Sigma p^i{}_1q^i{}_0\dfrac{p^i{}_1}{p^i{}_0}}{\Sigma p^i{}_1q^i{}_0}\cdot\frac{\Sigma p^j{}_0q^j{}_1\dfrac{p^j{}_0}{p^j{}_1}}{\Sigma p^j{}_0q^j{}_1}\lessgtr 1.$$

This inequality reduces to

$$0\lessgtr\Sigma\Pi\quad p^s{}_0p^r{}_1(p^{j_2}{}_0p^{i_2}{}_1q^i{}_0q^j{}_1)(p^i{}_0p^j{}_1-p^j{}_0p^i{}_1).$$

$$r\neq i,\ j$$

$$s\neq i,\ j$$

That there is no necessary fixed relationship would seem to be indicated by this formula. That this is the case is shown as follows:

For the set of values (β) we find $C=\frac{25}{24}>1$.

For the set of values

	p_0	q_0	p_1	q_1
(δ)	1	2	3	2
	1	3	2	1

we find $C=\frac{35}{36}<1$.

For the set of values

	p_0	q_0	p_1	q_1
(ϵ)	1	2	3	1.5
	1	3	2	1

we find $C=1$.

Using IV and applying the time reversal test we have to discover whether

$$D\equiv\frac{\Sigma p^i{}_1q^i{}_1\frac{p^i{}_1}{p^i{}_0}}{\Sigma p^i{}_1q^i{}_1}\cdot\frac{\Sigma p^j{}_0q^j{}_0\frac{p^j{}_0}{p^j{}_1}}{\Sigma p^j{}_0q^j{}_0}\gtreqless 1.$$

This inequality reduces to

$$0\gtreqless\Sigma\Pi\quad p^s{}_0p^r{}_1q^j{}_0q^i{}_1p^{j_2}{}_0p^{i_2}{}_1(p^i{}_0p^j{}_1-p^j{}_0p^i{}_1).$$
$$s\neq i,\ j$$
$$r\neq i,\ j$$

Whence it would seem that no necessary fixed relationship exists. That this is the case we illustrate by the following sets of values:

Thus for

p_0	q_0	p_1	q_1
2	2	3	1
2	3	2	4

we find $D=\frac{65}{66}<1$.

For the set of values

p_0	q_0	p_1	q_1
2	2	3	1
2	3	2	3

we find $D=\frac{91}{90}>1$.

Thus any statement that this index number has any fixed upward or downward bias is erroneous.

It is of interest to compare the relative size of index numbers computed on the basis of a weighted arithmetic average of relatives when I, II, III, IV are used as weights.

I vs. III: If an index number computed on the basis of a weighted arithmetic average of relatives is always greater when p_1q_0 are used as weights, than when p_0q_0 are used as weights, we have to show that

$$\frac{\Sigma p^i{}_0 q^i{}_0 \frac{p^i{}_1}{p^i{}_0}}{\Sigma p^i{}_0 q^i{}_0} < \frac{\Sigma p^j{}_1 q^j{}_0 \frac{p^j{}_1}{p^j{}_0}}{\Sigma p^j{}_1 q^j{}_0}.$$

This inequality reduces to

$$0 < \sum_{i=1}^{N} \sum_{j=1}^{N} q^i{}_0 q^j{}_0 [p^i{}_0 p^j{}_1 - p^j{}_0 p^i{}_1]^2 \cdot \prod_{s=1}^{N} p^s{}_0 (S \neq i, j).$$

Thus we see that I<III, with the single exception, when the percentage increase, or decrease, is the same for every commodity, then the two are equal.

II vs. IV: If an index number computed on the basis of a weighted arithmetic average of relatives is always greater when p_1q_1 are used as weights, than when p_0q_1 are used, then we have to show that

$$\frac{\Sigma p_0 q_1 \frac{p_1}{p_0}}{\Sigma p_0 q_1} \leq \frac{\Sigma p_1 q_1 \frac{p_1}{p_0}}{\Sigma p_1 q_1}.$$

This inequality reduces to

$$0 \leq \Sigma q^i{}_1 q^j{}_1 (p_0{}^i p^j{}_1 - p^j{}_0 p^i{}_1)^2 \prod_{s=1}^{N} p^s{}_0 (S \neq i, j).$$

This shows that II≤IV.

The equality holds only when the percentage increase (decrease) is the same for every commodity.

III vs. IV: We desire to discover whether

$$\frac{\Sigma p_1 q_0 \frac{p_1}{p_0}}{\Sigma p_1 q_0} \lessgtr \frac{\Sigma p_1 q_1 \frac{p_1}{p_0}}{\Sigma p_1 q_1}.$$

This inequality reduces to

$$0 \lessgtr \Sigma p^i{}_1 p^j{}_1 (p^i{}_0 p^j{}_1 - p^j{}_0 p^i{}_1)(q^i{}_0 q^j{}_1 - q^j{}_0 q^i{}_1) \prod_{s=1}^{N} p^s{}_0 (i \neq j,\ s \neq i, j).$$

Whence it would seem that there is no necessary fixed relationship. That this is the case is shown by the following:

For the set of values (β) we find

$$\text{III}=\frac{25}{10}, \qquad \text{IV}=\frac{26}{10}$$

whence $$\text{III}<\text{IV}.$$

For the set of values (α) we find

$$\text{III}=\frac{35}{14}, \qquad \text{IV}=\frac{34}{14}$$

Thus $$\text{III}>\text{IV}.$$

For the set of values (γ) we find $\text{III}=\text{IV}=\frac{5}{2}$.

II vs. III: We desire to discover whether

$$\frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \lessgtr \frac{\Sigma p_1 q_0 \dfrac{p_1}{p_0}}{\Sigma p_1 q_0}.$$

This inequality reduces to

$$0 \lessgtr \Sigma(p^i_0 p^j_1 - p^j_0 p^i_1)(p^i_0 p^j_1 q^j_0 q^i_1 - p^j_0 p^i_1 q^i_0 q^j_1) \prod_{s=1}^{N} p^s_0 (S \neq i, j).$$

Whence it would seem that there is no necessary fixed relationship. That this is the case is shown as follows:

For the set of values (β) we find $\text{II}=\text{III}$.
For the set (α) we find $\text{II}<\text{III}$.
For the set (δ) we find $\text{II}>\text{III}$.

I vs. II: We have to discover whether

$$\text{I} \equiv \frac{\Sigma p^i_1 q^j_0}{\Sigma p^i_0 q^i_0} \lessgtr \frac{\Sigma p^j_1 q^j_1}{\Sigma p^j_0 q^j_1} \equiv \text{II}.$$

This inequality reduces to

$$0 \lessgtr \Sigma\Sigma(q^i_0 q^j_1 - q^j_0 q^i_1)(p^i_0 p_1 - p^j_0 p^i_1)(i \neq j).$$

That no necessary fixed relationship exists can be shown as follows.
For the set of values (α) we find $\text{I}>\text{II}$.
For the set of values (β) we find $\text{I}<\text{II}$.
For the set of values (γ) we find $\text{I}=\text{II}$.

I vs. IV: That no necessary fixed relationship exists can be shown as follows. We have

$$\mathrm{I} \equiv \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}, \quad \mathrm{IV} \equiv \frac{\Sigma p_1 q_1 \frac{p_1}{p_0}}{\Sigma p_1 q_1}.$$

For the set of values (a) we find I<IV.

For the set of values

p_0	q_0	p_1	q_1
1	2	3	1
1	3	2	3

we find I>IV.

In the literature I find statements of which the following is typical: "A downward bias arises from weighting systems I and II, while an upward bias arises from weighting systems III and IV. This is in part capable of mathematical demonstration, and has in part been established by numerous trials."

In the light of what I have proved, it would seem time to omit the fallacies contained in the quotation above.

These statements are usually made without any qualification as to the type of index to which the weights are to be applied.

For the median of relatives, weighted by values in the base year, we have $\left(\frac{p_1}{p_0}\right)^{p_0 q_0}$. Applying the time reversal test we have to discover whether

$$M \equiv \left(\frac{p^i_1}{p^i_0}\right)^{p^i_0 q^i_0} \cdot \left(\frac{p^j_0}{p^j_1}\right)^{p^j_1 q^j_1} \gtreqless 1.$$

Obviously, it is an easy matter to find values of p_0, q_0, p_1, q_1 such that we might have M either less than or greater than unity. Hence again weighting by base year values may give an upward bias, a downward bias, or no bias.

Most statements made in the past regarding bias would seem to have been made on the basis of the numerical results obtained in a limited number of problems.

It is believed that we have given here for the first time definite mathematical formulas and proofs as to the existence or non-existence of bias, and, if bias exists, its character, for most of the cases considered.

The question under discussion in this paper is a purely mathematical one with respect to the existence or non-existence of bias whether due to type or weighting. There is another important question which is not touched in this paper. Namely, is there any relation between

price and quantity for a given commodity, and is there any relation between the price and quantity of one commodity and the price and quantity of another commodity, and if these relations exist are they strong enough to keep an index number computed in some given manner permanently greater than an index number computed in some other given manner? Furthermore, if for actual commodity prices the inequality is not always in the same direction what can be said definitely as to the proportion of cases for which the inequality is in a given direction? These questions I have not tried to answer in this paper. These are as yet unsolved problems.

THE POPULATION OF CHINA IN 1910[1]

By Walter F. Willcox, *Cornell University*

The population of China is the largest unknown part of the population of the earth. How far it is from determined appears in the fact that a recent and authoritative repository of information about that country gives three figures for the population at recent dates of the old eighteen provinces often called China Proper, the largest, 411,000,000, exceeding the smallest, 316,000,000, by 95,000,000, a number not much less than the population of all Latin America.[2] This it does without making any attempt to harmonize the figures or to indicate which is most trustworthy.

Western students of Chinese conditions, with a few exceptions, have followed the same course as the *China Year Book* and abandoned the problem as insoluble. It must be admitted that no final solution is in sight. But a question of this nature is mastered by reducing the range of possibilities and strengthening the evidence for an answer of a certain order. Along that road within a century progress has been clear and definite. In 1810 Malte-Brun wrote:

> Among the different accounts respecting China how shall we hit upon the true one? Has that country 27,000,000 inhabitants according to Sonnerat? or 55,000,000 according to the extract from the Official Gazette of Pekin? or 70,000,000 according to the Russians? or 100,000,000 as De Guignes believed? or 19,662,000 as Mr. Buesching informs us? or 200,000,000 agreeable to the statement of the missionaries? or, finally, 333,000,000 as a Chinese mandarin, perfectly veracious no doubt, assured Lord Macartney?[3]

The largest guess reported by Malte-Brun exceeded the smallest more than ten-fold; today no one places the population of China below 270,000,000 or above 510,000,000,[4] a difference of less than two-fold. Thus during the nineteenth century the range of the guesses was greatly reduced.

But progress in another direction has been more important. When

[1] This paper was presented before the Seventeenth Session of the International Statistical Institute, held at Cairo in the winter of 1927–28, and will be published in the Bulletin of the Institute. Any one interested in the development of the writer's opinions about the population of China will find them stated in "The Expansion of Europe," in *Studies in Philosophy and Psychology, A Commemorative Volume*, pp. 66–70, Houghton, Mifflin and Company (1906); "The Expansion of Europe in Population," in *American Economic Review*, Vol. V, pp. 742–4 (1915); "China's Population—400,000,000 or 300,000,000?" in *Chinese Students' Monthly*, Vol. XXII, pp. 23–9 (1926).

[2] *The China Year Book, 1926-1927*, p. 3.

[3] Malte-Brun, *Précis de la Géographie Universelle*, Paris, 1810. The quotation is from the translation published at Boston in 1824, Vol. 1, p. 549.

[4] These maximum and minimum figures are taken from the article on "Population" in Couling's *Encyclopædia Sinica*. They are reached by putting in one column the minimum estimate for each province whatever the source and in another column similarly the maximum estimate. No significant estimate for all China has been found as low as 270,000,000 or as high as 510,000,000, although the *Statesman's Year-Book for 1927* does refer to an estimate by an independent investigator (unnamed) of 545,000,000.

Malte-Brun wrote he could present nothing beyond a series of guesses mainly by western writers unsupported by evidence. All students of the subject today recognize the existence and importance of a long series of Chinese censuses, so-called, and the discussion turns upon the methods by which they were taken and the meaning, congruence and trustworthiness of their results.

There are two outstanding reasons for believing that another forward step towards determining the population of China is now possible. The first is that the enumeration of 1910 has furnished probably more trustworthy and certainly more abundant material for analysis than any earlier or later Chinese census. This was the opinion of Tenney, who wrote in 1911:

"The census of 1910 is probably the most accurate that has ever been taken [in China]," [1] and of Rockhill, who wrote a year later: "The first official Chinese attempt at a real enumeration of the population . . . seems entitled to more confidence than the enumerations which have heretofore appeared." [2]

A different opinion is entertained by Roxby who says that the China Continuation Committee's estimate of 1918–1919 and the Post Office Census of 1920 "represent the most scientific and probably the most reliable computation of the population of China that has yet been made." [3]

Examination of the China Continuation Committee's estimate gives ground for a different conclusion. Their table shows a population of 441,000,000 and that of the Post Office a population of 428,000,000, but the most definite statement in the text is that "Undoubtedly the exact population of China is considerably lower than most estimates now lead one to believe. . . . Perhaps the present population of the Chinese Republic lies somewhere between 350 and 400 millions." [4] This is in substance an admission that the official figures in their table are probably too large by between 10 and 20 per cent.

A second reason for believing that the problem can now be advanced towards a solution is that the more abundant and better material furnished by the census of 1910 has not yet been reviewed by any one trained in statistics and acquainted with the probabilities suggested by the figures of other Oriental countries like Japan, British India and the Philippines, in which statistics have advanced farther than in China. What is needed is to bring both familiarity with Chinese conditions and familiarity with statistical methods and results to bear upon

[1] *U. S. Daily Consular and Trade Reports* for July 13, 1911, pp. 184–7.

[2] W. W. Rockhill, "The 1910 Census of the Population of China" in *T'oung-Pao*, Vol. XIII (1912), p. 117. Reprinted in Am. Geog. Society Bulletin Vol. XLIV (1912): pages 668–673.

[3] "Distribution of Population in China" in *Geographic Review*, 15 (1925), pp. 1–24.

[4] *The Christian Occupation of China*, p. 11.

the problem. The writer has tried to supplement his deficiencies in the former field by correspondence and conference with men long resident in China.[1]

The history of census-taking in the Celestial Empire falls into two periods separated by the imperial edict of A.D. 1712 which decreed "that the number of families given in the enumeration of the preceding year should remain the invariable basis for assessment of the crown taxes." [2]

The statement of Sacharoff is:

> In A.D. 1712 an edict was issued which decreed that the number of tax-paying people in A.D. 1711 should forever remain the standard of those subject to crown service; the rest of the population, whatever might be its increase, being exempted from the payment of taxes.[3]

Parker puts it a little differently:

> The poll-tax was merged [sc. in 1712] in the land-tax. Each taxable unit, say, was an acre of first-class ground and there were at that time about 100,000,000 English acres taxed. . . . As the revenue was sufficient the Emperor decided that in future, no matter how the population might increase, the land being a fixture ought never to pay more than 100,000,000 units.[4]

The result of the radical change was to cut off the Chinese census from any use in administration. Before 1712 imperial taxation was based on it at least in part; subsequently for two centuries, 1712–1910, nothing of any local interest was based on it. The census was maintained "in order to present the rolls of the census to heaven and to hand the same over to the college of the imperial historians." [5]

The effect of this change upon the character and trustworthiness of the Chinese censuses taken between 1712 and 1910 is not of importance for the present argument, which aims only to show that the census of 1910, like its predecessors before 1712, had a practical purpose. The purpose was different, to be sure; the new census sought to give a basis not for imperial taxation but for apportioning seats in the newly established provincial assemblies. On this point Rockhill, who was in 1908 American Minister at Peking and for many years had interested himself in the question of China's population, wrote:

> When the Empress Dowager took the momentous step of promising the people a representative government, the need of a census became at once felt as it probably never had been before. . . . In an Imperial Edict of August 27,

[1] Among those who have helped him are Professor E. T. Williams of the University of California, Professor Paul Langhans, Editor of *Petermanns Mitteilungen*, Professor Erich Hauer of the University of Berlin, Professor Hu Shih of the University of Peking, Mr. Sao Ke Alfred Sze, Chinese Minister to the United States, the late W. W. Rockhill, American Minister to China, 1905–1909, and most of all Mr. Raymond P. Tenney of the American Department of State.

[2] W. W. Rockhill, "Inquiry into the Population of China," in *Smithsonian Report for 1904*, page 663.

[3] Sacharoff, *Rise and Fall of the Chinese Population* (1864), p. 43.

[4] E. H. Parker, *China, Past and Present* (1903), p. 26.

[5] Sacharoff, *idem*.

> 1908 . . . it was prescribed that in 1909 the Ministry of the Interior with the assistance of the Viceroys and Governors was to begin taking a census by households . . . and report the results to Peking in 1910. In 1911 a further census (by heads) was to be made and reported to the Government in 1912.

This edict laid down a nine-year program of reforms introducing constitutional government. In the first year, 1908–1909, among other steps towards that end census regulations were to be issued; in the second year, 1909–1910, the census was to be taken throughout the Empire; in the third year, 1910–1911, the Viceroys and Governors were to report on the results of the census in their provinces; in the fourth year, 1911–1912, there was to be an investigation of the census.[1] Evidence seems to indicate that during the first three years the program was carried out. The census was taken in 1910 and a year later its results were reported. They are found in a Report to the Throne made on February 27, 1911, by the Chinese Ministry of the Interior and published in the Peking Government Gazette.

After not a little search, including two fruitless examinations of the Chinese material in the Library of Congress, a copy of the Chinese text of this Report has been found and also in typewritten form the text of an English translation.[2] Table I reproduces the translation of the summary of the earlier part of the Chinese census of 1910, that upon households, from Mr. Tenney's translation with a few changes which his own review of the subject or our discussions have led him or me to make. All figures in the columns headed "Previous Numbers" and "Total" and the three totals at the end are additions to the original. Figures in the last column are the sums of those in the two preceding. Those in the first column require a word of explanation. The interpretation put on the Chinese Report by all who have studied it is that the districts to which the figures relate are mutually exclusive (for example, the district "Peking Eight Banners" is not included in the first district "Peking"). Under that interpretation the number of households in China is to be determined by adding the numbers reported from each district. To get the number of households in Chihli province, for example, it is necessary to add the number in all enumeration districts lying within that province. Every such district has been given the number 1 in the first column of Table I and in the same way every other

[1] *China Year Book 1912*, Chapter XXI, *passim*.

[2] The former is in the Library of Congress (filed under Orientalia Chinese *B* 312 26 and in *Cheng Chih Kuan Pao* dated Hsuan T'ung, 3rd year, 1st moon, 29th day); the latter is in the files of the American Department of State. I owe their discovery and help in their interpretation to Mr. Raymond P. Tenney, author of the translation and now attached to the Division of Far Eastern Affairs in the Department of State. The copy of the Report in the Library of Congress is the only one which has been unearthed as a result of inquiries addressed also to several other American libraries, the libraries of the British Museum, of the Royal Statistical Society and of the Oriental Seminar at Berlin, to Chinese diplomatic representatives at Washington, London and Berlin and to the American Minister at Peking.

province has been given a number determined by its geographical position according to a common arrangement of the eighteen provinces shown in Table II. The numbers greater than 18 represent parts of China outside of the old eighteen provinces, 19 is Sinkiang or Eastern Turkestan, 20–22 inclusive are Manchuria, and 23 is Mongolia.

TABLE I

NUMBER OF HOUSEHOLDS IN THE CHINESE EMPIRE IN 1910

Province number	Enumeration district	Householders	Renters	Total
1	Peking	68,561	70,009	138,570
1	Shun T'ien Fu (exclusive of Peking)	600,797	91,899	692,696
20	Fengtien (28 districts)	549,910	249,926	799,836
21	Kirin (whole province)	422,781	316,680	739,461
22	Heilungkiang (whole province)	145,929	95,082	241,011
1	Chihli (whole province)	3,606,936	557,153	4,164,089
	Kiangsu (under Nanking, larger districts)	807,909	170,097	.....
	(under Soochow, all)	1,697,499	472,629	.....
6	Total, Kiangsu	.....	.....	3,148,134
7	Anhui (whole province)	2,486,896	654,288	3,141,184
2	Shantung (whole province)	5,143,699	234,173	5,377,872
3	Shansi (whole province)	1,520,031	470,004	1,990,035
8	Honan (whole province)	3,969,308	692,258	4,661,566
4	Shensi (whole province)	1,319,210	282,234	1,601,444
5	Kansu (whole province)	711,000	195,639	906,639
19	Sinkiang (whole province)	385,845	67,632	453,477
12	Fukien (whole province)	1,699,067	677,788	2,376,855
11	Chekiang (whole province)	2,524,635	1,363,677	3,888,312
13	Kiangsi (bulk of province)	2,287,421	1,098,907	.....
	Treaty ports	20,509	16,316	.....
	Boat families:			
	provincials	9,023	.....	.....
	extra provincials	7,697	.....	.....
	Total, Kiangsi	.....	.....	3,439,873
9	Hupeh (whole province)	3,783,179	749,352	4,532,531
14	Hunan (whole province)	2,574,128	1,714,036	4,288,164
10	Szechwan (55 districts)	2,321,725	937,992	3,259,717
18	Kwangtung (whole province)	4,358,473	683,307	5,041,780
17	Kwangsi (whole province)	1,097,539	77,005	1,174,544
16	Yunnan (whole province)	1,328,292	219,722	1,548,014
15	Kweichow (whole province)	1,634,782	136,751	1,771,533
1	Peking Eight Banners	118,783	.....	118,783
1	Imperial Household	4,571	.....	4,571
1	Peking Suburban Camps	56,536	17,656	74,192
1	Peking Gendarmerie:			
	Left Wing, 4 wards	486	368	.....
	Right Wing, 5 wards	538	240	1,632
1	Eastern Tombs Banner Camps	2,981	1,225	4,206
1	Western Tombs Banner Camp	908	138	1,046
1	Ma Lan Chen jurisdiction	586	340	926
1	T'ai Ning Chen jurisdiction	2,209	766	2,975
23	Jehol Mongol Banners	54,994	2,764	57,758
23	Ch'a-har jurisdiction	12,938	.....	12,938
1	Mi Yun Manchu garrison	1,917	.....	1,917
1	Shanhaikwan Manchu garrison	1,949	.....	1,949
2	Ch'ing Chou Manchu garrison	2,405	.....	2,405
3	Sui Yuan Ch'eng Manchu garrison	2,765	.....	2,765
4	Hsi An Manchu garrison	2,525	1,373	3,898
5	Liang Chou Manchu garrison	794	.....	794
19	Ili Manchu garrison	13,214	.....	13,214
12	Foochow Manchu garrison	1,738	546	2,284
10	Ch'engtu Manchu garrison	2,516	1,341	3,857
18	Canton Manchu garrison	6,885	3,753	10,638
23	Uliasutai jurisdiction	13,516	.....	13,516
23	Tarbagatai jurisdiction	3,887	.....	3,887
23	Cobdo jurisdiction	17,108	.....	17,108
5	Hsining jurisdiction	1,221	811	2,032
23	Urga jurisdiction	40,105	.....	40,105
½ to 10 ½ to 16	Szechwan-Yunnan Border Lands jurisdiction	46,362	2,512	48,874
	Total	47,497,218	12,328,389	59,825,607

Table I furnishes the data for determining the number of households in seventeen of the eighteen provinces of China Proper. In the case of Szechwan the returns are confessedly incomplete, only 55 of the 175 districts or counties reporting. The revolution which dethroned the Manchu dynasty in the following year started in Szechwan and at the time of the census of 1910 parts of that remote province were in turmoil. It has been assumed that the average number of households per district in the missing 120 districts was the same as in the 55 which did report. On this assumption there were 10,394,194 households in Szechwan, a number which, however far from the truth, is doubtless much nearer to it than the 3,288,011 households reported.

There is a question whether the last enumeration district in Table I lay within China Proper or in Tibet. It has been assigned to the former, because to do so increases the number of households and so the estimated population of China. In default of special information it has been divided equally between Szechwan and Yunnan.

With these explanations it is easy to see how Table I has been interpreted into Table II.

TABLE II

NUMBER OF HOUSEHOLDS IN CHINA PROPER BY THE CHINESE CENSUS OF 1910

Province	Number of households
North China	15,095,436
Chihli	5,207,552
Shantung	5,380,277
Shansi	1,992,800
Shensi	1,605,342
Kansu	909,465
Central China	25,877,609
Kiangsu	3,148,134
Anhui	3,141,184
Honan	4,661,566
Hupeh	4,532,531
Szechwan	10,394,194
South China	23,566,434
Chekiang	3,888,312
Fukien	2,379,139
Kiangsi	3,439,873
Hunan	4,288,164
Kweichow	1,771,533
Yunnan	1,572,451
Kwangsi	1,174,544
Kwangtung	5,052,418
China Proper	64,539,479

The next and more difficult problem is to estimate the number of persons in these 64,540,000 households. For that purpose it is necessary to determine as well as may be the average number of persons to a household. The census of 1910 was more important than any other Chinese census ancient or modern because it was the only one, so far as I know, which attempted to determine the population as well as the number of households. But when examining the evidence it presents

it will be well to have in mind what was known before 1910 about the average size of a Chinese household. The following passage in Rockhill's paper of 1904 summarizes the best knowledge or opinion about that question:

> There is much uncertainty as to the number of individuals contained in each recorded "household" or hu, . . . In the time of Mencius (fourth century B.C.) the "family" (chia) was supposed to comprise eight mouths (k'ou). This was the number of persons whom 100 mou (about 15 acres) of medium land were computed to support. Under some dynasties (as the Han) it would seem that the "household" comprised from 4.8 to 5.2 individuals; in others, the T'ang, for example, it rose to 5.8. During others, as the Sung, it was only a fraction over 2 persons, according to Sacharoff, though Biot contends that in this period it was a fraction more than 5 persons, as in the preceding period of the T'ang. Under the Yuan dynasty, according to Amiot, the "household" comprised 5 persons, and in the succeeding Ming dynasty it seems to have varied from about 5 to over 6.6. Even during the present dynasty we are in grave doubt as to the numeric value of the term hu ("household," "family"). Father Amiot and other foreign writers have thought it represented 5 persons, de Guignes says 2 to 3, but in the opinion of E. H. Parker it averaged 6 persons. In the census of 1842, which gave the number of households and of individuals, the former averaged 2.3 persons to the family; and in a census of the city of Peking for 1846 it averaged 3.1.[1] I am disposed to accept 4 as a fair figure for enumerations of the eighteenth and nineteenth centuries.[2]

On this passage several comments may be made:

1. The most acute and best informed of recent writers on the population of China believed in 1904 that in modern times the average number of persons to a Chinese family or household (both English words are used as a translation of one Chinese term) was about four. Rockhill's change of opinion in 1912 and the reasons for it will be discussed later.

2. The reasons assigned for this opinion do not furnish it with any statistical support. The figure is not even an estimate, but only the guess of a well-informed man.

3. That the average number of persons to a household at different times under the Manchu dynasty (1644–1912 A.D.) varied between two and six persons, as the passage perhaps implies, is incredible. The average number of persons to a household in the United States was 5.7 in 1790 and decreased slowly by about one-tenth of a person in each decade to 4.3 persons in 1920. Probably the change in China if any during the same period was much less than that in the United States.

4. A knowledge of statistical probabilities and especially of the results in other Oriental countries where statistics have been more developed is almost as necessary as a knowledge of Chinese conditions if one is to escape error.

[1] These statements I have been unable to verify.

[2] W. W. Rockhill: "An Inquiry into the Population of China" in *Smithsonian Report for 1904*, pp. 659 ff.

After this survey of what was known at the opening of the present century about the average size of a Chinese household the new evidence to be derived from a further analysis of the census of 1910 may be examined.

This evidence is found in the second part of the Report to the Throne summarizing fragmentary returns of the census of population. They relate to sixteen districts lying in six of the eighteen provinces well enough distributed to be representative of the whole. The report covers the whole population of four provinces, 89 districts in Shansi and 125 in Szechwan. Its results are brought together by provinces in Table III.

TABLE III

RESULTS OF THE CENSUS BY INDIVIDUALS, FOR CERTAIN CHINESE PROVINCES, 1910

Province	Males	Females	School children	Able-bodied males
Chihli	14,056,900	11,650,857	2,194,496	4,747,653
Shansi (89 districts)	4,528,445	3,400,719	493,707	1,587,191
Szechwan (125 districts)	7,133,090	5,301,980	1,338,330	2,595,479
Chekiang	7,004,082	5,909,237	1,030,336	3,057,912
Kiangsi	8,237,379	6,235,300	.....	.....
Kweichow	4,636,965	3,866,998	862,951	1,987,836
Total	45,596,861	36,365,091	5,919,820	13,976,071

The interpretation of Table III is more difficult than that of Table I. First there are uncertainties of translation. Rockhill interprets the headings of the first and second columns as "Adult Males" and "Adult Females." I am informed that the meaning of the Chinese term is not limited to adults and suspect that Rockhill was misled by the third column "School-children" into assuming that only adults had been counted in the first two classes. Between columns 2 and 3 is a Chinese term which Tenney at first translated by "Additional Enrollment" for which he would now substitute "Additional Investigation." This term may apply equally to columns 3 and 4 and that application, although not necessary, is perhaps to be preferred. Its introduction seems to support an interpretation which I had come to accept before this Chinese word was called to my attention. It seems likely that the whole population was to be reported in two great classes of males and females and in addition two special classes, namely, children in school or of school age (6 to 16 years of age Rockhill interprets it) and the males fit for military service or able-bodied males ("probably of from 16 to 45 years of age" Rockhill interprets it). If the instructions to the provincial authorities, one mention of which I have found, could be discovered they might determine this crucial point.

The main problem in the interpretation of Table III is this, "How is the total population to be determined from the figures in these four columns?" My answer is, "By adding the figures in the first two columns." Before examining that interpretation it will be well to consider two others. In his original report Mr. Tenney added the four columns and by that path arrived at 5.5 for the average size of a Chinese household, a figure which since then has been widely accepted. He now, I understand, would omit at least the figures in the fourth column on the ground that they are probably included in the first.

Rockhill added the figures in the first three columns to get the total population and thus reached 4.8 as the average size of a family. But this interpretation is vitiated in my opinion by his misunderstanding the figures in the first two columns as confined to adults and by an oversight of the word "additional investigation" or a failure to read into it what seems to me its natural implication. The Chinese text, then, favors but does not require interpreting the first two columns as including persons of all ages. In that case their sum would be the total population.

What light does internal evidence throw upon Table III? It will be noticed that the males outnumber the females by more than 9,000,000. This result seems improbable; certainly, it is unparalleled. The proportion of males reported in these six provinces is about equal to that among the foreign-born population of the United States or among the total population of ranching states like Montana or Wyoming. What is known about the approximate equality in the number of the sexes at birth and about the little influence of migration in disturbing that initial equality in great masses of population compels one to look with doubt upon these returns. If they are correct they point to an amount of female infanticide in China heretofore hardly suspected. If they are incorrect the error probably lies in the wholesale omission of females. In favor of their correctness it must be remembered that every Oriental country in which the population has been counted with distinction of sex shows an excess, though a far smaller excess, of males.

This difficulty may be met by making one or the other of two assumptions, first, that the figures are correct and, secondly, that many females were omitted and that their true proportion in these Chinese provinces was the same as that in British India, the Oriental country apart from China with the greatest proportion of males. The second assumption would cut down the excess of males in these provinces from over 9,000,000 to about 2,500,000. Very likely both female infanticide on a great scale and the omission of many females from the count concurred to bring about the result. If so the truth would lie between the two ex-

tremes. Of the two assumptions I prefer the former, namely, that the figures are correct.

Does internal evidence throw any light upon Rockhill's interpretation that the first two columns are confined to adults and that to get the total population not only the children of school age (6 to 16 years)[1] in column 3 but the children under 6 must be added?

Of these four columns the one whose meaning is most certain is probably the last. Three interpreters working independently agree that this column gives the number of males able to bear arms.[2] The age limits of the group are not easy to define; Rockhill suggests "probably of from 16 to 45 years of age." It is my belief that he puts the lower age limit too early. In the United States the militia age is 18–44 and in the leading European countries before the War, as in Japan, service with the colors began at 20. So I will assume that column 4 includes all males between 20 and 44 years of age or, what amounts for present purposes to the same thing, that the number of males within that age period who were not reported (perhaps because they were physically unfit to bear arms) was balanced by younger males who were. On this assumption the following per cents for several Oriental countries and the United States have been computed to compare with the Chinese per cent.

Country	Date	Per cent that males 20–44 years of age in certain countries make of		
		All males	Males at least 6 years of age	Males at least 20 years of age
Japan	1918	34.7	41.0	62.8
British India	1921	36.8	43.1	69.5
Philippines	1903	34.6	42.2	68.9
Australia	1921	37.6	43.3	62.5
Egypt	1917	33.6	40.2	66.7
United States	1920	38.1	43.3	62.5

The ratio of the number of Chinese able to bear arms (Table III, Column 4) to the number in Column 1 is 37.4 per cent. It is clear that this ratio belongs in the first rather than in the second or third column above. If so and if the men able to bear arms included all between 20 and 44 years of age then the male population in column 1 includes

[1] If the figure in the first column includes all males and if the proportion of the sexes in these provinces is about the same as in British India, then the school children in column 3 are little more than one-third of the actual number of children between 6 and 16 as estimated from their proportion in other Oriental countries. Any other assumption about columns 1 and 2 would increase the apparent shortage in column 3.

[2] They are Tenney, Rockhill and Hauer, whose translation of the "Report to the Throne" into German was made in Berlin for the German Embassy and printed in *Petermanns Mitteilungen*, Vol. 57, part 2 (1911), pp. 255 ff.

males of all ages (even those less than six years of age) and not merely adult males or adult males and children of school age.

Both internal and external evidence thus concur in discrediting the hypothesis that columns 1 and 2 include only adults or only those over six years of age. It follows that the total population of these provinces is probably the sum of the first two columns. If this interpretation be accepted for these five provinces at least and presumably therefore for the others, the average size of a Chinese household is not 5.5 persons or 4.8 persons but 4.3 persons. If many females were omitted from the count and if in fact the proportion of females to males in China was the same as in British India then the average size of a household was 4.5 persons. The former hypothesis seems preferable but in Table IV the population of China Proper has been computed on both.

TABLE IV

POPULATION OF CHINA PROPER OR THE EIGHTEEN PROVINCES BY THE CENSUS OF 1910

Province	Number of households	On the basis of	
		4.3 persons to a household	4.5 persons to a household
North China	15,095,436	64,910,000	67,930,000
Chihli	5,207,552	22,392,000	23,434,000
Shantung	5,380,277	23,135,000	24,211,000
Shansi	1,992,800	8,569,000	8,968,000
Shensi	1,605,342	6,903,000	7,224,000
Kansu	909,465	3,911,000	4,093,000
Central China	25,877,609	111,274,000	116,449,000
Kiangsu	3,148,134	13,537,000	14,167,000
Anhui	3,141,184	13,507,000	14,135,000
Honan	4,661,566	20,045,000	20,977,000
Hupeh	4,532,531	19,490,000	20,396,000
Szechwan	10,394,194	44,695,000	46,774,000
South China	23,566,434	101,336,000	106,049,000
Chekiang	3,888,312	16,720,000	17,497,000
Fukien	2,379,139	10,230,000	10,706,000
Kiangsi	3,439,873	14,791,000	15,479,000
Hunan	4,288,164	18,439,000	19,297,000
Kweichow	1,771,533	7,618,000	7,972,000
Yunnan	1,572,451	6,762,000	7,076,000
Kwangsi	1,174,544	5,051,000	5,286,000
Kwangtung	5,052,418	21,725,000	22,736,000
China Proper	64,539,479	277,520,000	290,428,000

In Table IV the most doubtful figures are probably those for Szechwan. It will be remembered that the number of households reported from that province was multiplied by 3.18 because the total number of its districts was 3.18 times the number from which reports were received. In his article published in 1904 Rockhill attempted to estimate the population of Szechwan. After rejecting the estimate of 71,000,000 made by the Board of Revenue in 1885 with the comment

"All foreign writers agree that it is quite impossible to believe that any such population exists—or can exist in it," he gave the following estimates:

Source	Date	Estimate
Maritime Customs at Ch'ung-K'ing	1891	30,000,000—35,000,000
Lyons Commercial Mission	1895–6	40,000,000—45,000,000
G. J. L. Litton	1898	43,000,000
F. A. S. Bourne	1898	45,000,000—55,000,000
Hosie	1904	45,000,000

To these should be added the figure in his article on the 1910 census, 54,500,000. This is based on a report of the Viceroy of Szechwan upon the census of 1910 printed in a Chinese newspaper. The average of these six estimates is 43,000,000–46,000,000 which agrees pretty well with the estimates in Table IV.

The foregoing analysis of the evidence seems to justify the conclusion that the population of China Proper in 1910 was probably between 275,000,000 and 300,000,000 and nearer the smaller than the larger number. Perhaps 280,000,000 is as close an approximation as the evidence permits.

The ablest writer on Chinese population is Rockhill. The general result of the preceding discussion is to reduce his estimate of 311,000,000 for the population of China Proper in 1910 (accepting his very doubtful figure of 9,000,000 for the omitted children under six years of age) to about 280,000,000 and so to bring the result into line with his earlier conclusion. Then he wrote: "The conviction is forced on me that the present population of China Proper does not exceed 275,000,000 and is probably considerably under this figure." The foregoing argument places his earlier result on a sounder statistical foundation than it has hitherto had.

To the figures in Table IV estimates may be added for the outlying districts of Sinkiang, Manchuria, Mongolia and Tibet and thus an approximate figure reached for the population of all China. For Sinkiang and Manchuria some material is available in the published returns of the census of 1910. From Fengtien in Manchuria returns were received for only 28 districts. As it now has twice as many, the number of households for that province in Table I has been doubled. For Tibet no information was published and for Mongolia the fragmentary material gives no basis for an estimate. For these districts, therefore, Rockhill's estimates of 1912 have been accepted. Thus Table V has been constructed.

TABLE V

ESTIMATED POPULATION OF OUTLYING DISTRICTS OF CHINA, 1910

Outlying districts	Number of households	Population on the basis of	
		4.3 persons to a household	4.5 persons to a household
Sinkiang	466,691	2,007,000	2,100,000
Manchuria	2,580,144	10,837,000	11,587,000
Fengtien	1,599,672	6,719,000	7,199,000
Kirin	739,461	3,106,000	3,328,000
Heilungkiang	241,011	1,012,000	1,060,000
Mongolia	.....	1,800,000	1,800,000
Tibet	.....	2,000,000	2,000,000
Total Outlying Districts	.....	16,644,000	17,487,000
China Proper	.....	277,520,000	.....
All China	.....	294,164,000	.....

The population of the Republic of China with its dependencies is not far from 295,000,000.

The foregoing estimate has already been compared with Rockhill's; it should be compared also with the best European results. The only authors familiar with geography and statistics who have examined the population of China over a series of years with penetrating analysis and trained critical ability are Behm, Wagner and Supan.[1] The general result of their discussions has been to estimate the population of China much below current official figures and the further their analysis was pushed the lower their results became. At the beginning in 1874 Behm based himself on the Chinese census of 1842 and estimated the population of China at 405,000,000. In 1882 he accepted the Chinese census of 1812 as probably more trustworthy than that of 1842 and in consequence reduced his estimate to 350,000,000. In 1901 Supan discussed a padding of the Chinese figures in 1775 by about 48,000,000 which had remained in all subsequent figures and to which Sacharoff had previously called attention. Supan devised an ingenious method for squeezing out this water and as a result reduced his estimate to 320,000,000. But this was long before the evidence derivable from the Chinese census of 1910 to which the present paper is confined had become available.

[1] In "Die Bevölkerung der Erde" Numbers I–XII (1872–1904).

THE STANDARD DEVIATION OF THE CORRELATION COEFFICIENT

By A. L. Bowley, *University of London*

The standard deviation of the correlation coefficient,

$$r=\frac{1}{n}\sum_{t=1}^{t=n}x_t y_t \div \sigma_x\sigma_y,$$

where there are n pairs of variables, such as x_t, y_t, measured from their averages, is generally given as $\frac{1-r^2}{\sqrt{n}}$. This was only proved to be correct when the joint frequency surface of x, y is normal in Professor Karl Pearson's original study; but Professor Edgeworth showed that an artificial normality could be produced by breaking up very numerous observations into batches,[1] so that the same standard deviation was approximately obtained if the x, y surface was moderately abnormal.

The full form of this standard deviation as given by Dr. Sheppard is

$$\sigma^2_r=\frac{r^2}{n}\left\{\frac{M_{22}}{r^2\sigma^2_x\sigma^2_y}+\frac{M_{40}}{4\sigma^4_x}+\frac{M_{04}}{4\sigma^4_y}-\frac{M_{31}}{r\sigma^3_x\sigma_y}-\frac{M_{13}}{r\sigma_x\sigma^3_y}+\frac{M_{22}}{2\sigma^2_x\sigma^2_y}\right\},$$

where $M_{st}=$mean of $x^s y^t$, $=|x^s y^t|$ in the notation used in the sequel, for all values of s and t.

In the normal correlation surface $M_{40}=3\sigma^4_x$, $M_{04}=3\sigma^4_y$, $M_{22}=(2r^2+1)\sigma^2_x\sigma^2_y$, $M_{31}=3r\sigma^3_x\sigma_y$, and $M_{13}=3r\sigma_x\sigma^3_y$, so that

$$\sigma^2_r=\frac{r^2}{n}\left\{\left(2+\frac{1}{r^2}\right)+\frac{3}{4}+\frac{3}{4}-3-3+r^2+\frac{1}{2}\right\}=\frac{1}{n}(r^4-2r^2+1)=\frac{(1-r^2)^2}{n},$$

the form commonly used.

It is proposed to examine the relationship between these formulae when normality is not assumed.

Let x and y be each the sum of m variables,

$$x=u_1+\ \ldots\ +u_t+\ \ldots\ u_m$$
$$y=v_1+\ \ldots\ +v_t+\ \ldots\ v_m,$$

all quantities being measured from their averages, such that any pair u_t, v_t may be related, but all other terms such as u_s, u_t . . . u_s, v_t . . . v_s, v_t are independent.

[1] *Journal of the Royal Statistical Society*, 1908. pp. 510–1.

Suppose that the frequencies of the u's, v's and pairs u_t, v_t differ from normality to an extent measured by d, δ and ϵ in the following equations; where ${}_t\sigma_u$, ${}_t\sigma_v$ are the standard deviations of u_t and v_t, and ρ_t is their correlation coefficient.

$$\begin{aligned}
&|u^4{}_t| = (3+{}_t\epsilon_u){}_t\sigma^4{}_u, \quad |v^4{}_t| = (3+{}_t\epsilon_v){}_t\sigma^4{}_v,\\
&|u^3{}_t v_t| = (3\rho_t+{}_td_u){}_t\sigma^3{}_u \cdot {}_t\sigma_v, \quad |u_t v^3{}_t| = (3\rho_t+{}_td_v){}_t\sigma_u \cdot {}_t\sigma^3{}_v,\\
&|u^2{}_t v^2{}_t| = (2\rho^2{}_t+1+\delta_t){}_t\sigma^2{}_u \cdot {}_t\sigma^2{}_v.
\end{aligned}$$

We have from the independence assumed, such relationships as

$$|u^2{}_t \cdot v^2{}_s| = |u^2{}_t||v^2{}_s| = {}_t\sigma^2 u \cdot {}_s\sigma^2{}_v, \ |u_t v_s| = |u_t||v_s| = 0, |u^3{}_t v_s| = |u^3{}_t||v_s| = 0,$$

and in the sequel zero terms are omitted.

$$\begin{aligned}
\sigma^2{}_x &= |u_1+u_2+ \ \ldots \ |^2 = \Sigma_t\sigma^2{}_u = m\sigma^2{}_u, \text{ say},\\
\sigma^2{}_y &= \Sigma_t\sigma^2{}_v = m\sigma^2{}_v, \text{ say},\\
r\sigma_x\sigma_y &= |xy| = |(u_1+u_2+ \ \ldots \ )(v_1+v_2+ \ \ldots \ )| \Sigma u_t v_t = \Sigma\rho_t \cdot {}_t\sigma_u \cdot {}_t\sigma_v\\
M_{40} &= |x^4| = |(u_1+u_2+ \ \ldots \ )^4| = |u^4{}_1+ \ \ldots \ +6u^2{}_1u^2{}_2+ \ \ldots \ |\\
&= 3(\Sigma|u^2{}_t|)^2+\Sigma|u^4{}_t|-3\Sigma|u^2{}_t|^2 = 3(\Sigma_t\sigma^2{}_u)^2+\Sigma\{|u^4{}_t|-3{}_t\sigma^4{}_u\}\\
&= 3\sigma^4{}_x+\Sigma_t\epsilon_u \cdot {}_t\sigma^4{}_u.
\end{aligned}$$

Write $\Sigma_t\epsilon_u \cdot {}_t\sigma^4{}_u = \sigma^4{}_u\Sigma_t\epsilon_u \cdot \left(\frac{{}_t\sigma_u}{\sigma_u}\right)^4 = m\sigma^4{}_u\epsilon_u$, so that ϵ_u is a weighted average of the ${}_t\epsilon_u$'s.

Then $M_{40} = 3\sigma^4{}_x+m\sigma^4{}_u \cdot \epsilon_u$.

Similarly $M_{04} = |y^4| = 3\sigma^4{}_y+m\sigma^4{}_v \cdot \epsilon_v$, where ϵ_v is a weighted average of the ${}_t\epsilon_v$'s.

$$\begin{aligned}
M_{31} = |x^3y| &= |(u_1+u_2+ \ \ldots \ )^3(v_1+v_2+ \ \ldots \ )|\\
&= |(u^3{}_1+ \ \ldots \ +3u^2{}_1u_2+3u_1u^2{}_2+ \ \ldots \ )(v_1+v_2+ \ \ldots \ )|\\
&= \Sigma|u^3{}_t v_t|+3\Sigma|u^2{}_s \cdot u_t v_t|\\
&= \Sigma(3\rho_t+{}_td_u){}_t\sigma^3{}_u \cdot {}_t\sigma_v+3\Sigma_s\sigma^2{}_u \cdot \rho_t \cdot {}_t\sigma_u \cdot {}_t\sigma_v\\
&= 3\Sigma\{\rho_t \cdot {}_t\sigma_u \cdot {}_t\sigma_v\Sigma_t\sigma^2{}_u\}+\Sigma_t d_u \cdot {}_t\sigma^3{}_u \cdot {}_t\sigma_v\\
&= 3\sigma^2{}_x\Sigma\rho_t \cdot {}_t\sigma_u \cdot {}_t\sigma_v+\Sigma_t d_u \cdot {}_t\sigma^3{}_u \cdot {}_t\sigma_v\\
&= 3r\sigma^3{}_x\sigma_y+m\sigma^3{}_u\sigma_v d_u,
\end{aligned}$$

where $\Sigma_t d_u \cdot {}_t\sigma^3{}_u \cdot {}_t\sigma_v = \sigma^3{}_u\sigma_v\Sigma_t d_u\left(\frac{{}_t\sigma_u}{\sigma_u}\right)^3 \cdot \frac{{}_t\sigma_v}{\sigma_v} = \sigma^3{}_u\sigma_v \cdot md_u$ and d_u is a weighted average of the ${}_td_u$'s.

Similarly $M_{13} = |xy^3| = 3r\sigma_x\sigma^3{}_y+m\sigma_u\sigma^3{}_v d_v$, where d_v is a weighted average of the ${}_td_v$'s.

$$\begin{aligned}
M_{22} = |x^2y^2| &= |(u_1+u_2+ \ \ldots \ )^2(v_1+v_2+ \ \ldots \ )^2|\\
&= |(u^2{}_1+ \ \ldots +2u_1u_2+ \ \ldots \ )(v^2{}_1+ \ \ldots \ +2v_1v_2+ \ \ldots \ )|\\
&= \Sigma|u^2{}_t v^2{}_t|+\Sigma|u^2{}_s v^2{}_s|+4\Sigma|u_1u_2v_1v_2|\\
&= \Sigma(2\rho^2{}_t+1+\delta_t){}_t\sigma^2{}_u \cdot {}_t\sigma^2{}_v+\Sigma|u^2{}_s||v^2{}_t|+4\Sigma|u_s v_s||u_t v_t|\\
&= \Sigma(2\rho^2{}_t+1){}_t\sigma^2{}_u \cdot {}_t\sigma^2{}_v+\Sigma_s\sigma^2{}_u \cdot {}_t\sigma^2{}_v+4\Sigma\rho_s\rho_t \cdot {}_s\sigma_u \cdot {}_s\sigma_v \cdot {}_t\sigma_u \cdot {}_t\sigma_v\\
&\qquad +\Sigma\delta_t \cdot {}_t\sigma^2{}_u \cdot {}_t\sigma^2{}_v
\end{aligned}$$

$$= \Sigma_t \sigma^2_u \cdot \Sigma_t \sigma^2_v + 2(\Sigma \rho_t \cdot {}_t\sigma_u \cdot {}_t\sigma_v)^2 + \Sigma \delta_t \cdot {}_t\sigma^2_u \cdot {}_t\sigma^2_v$$
$$= \sigma^2_x \sigma^2_y (1+2r^2) + m\delta_t \cdot \sigma^2_u \sigma^2_v$$

where $m\delta = \sigma^2_u \sigma^2_v \Sigma \delta_t \left(\frac{{}_t\sigma_u}{\sigma_u}\right)^2 \left(\frac{{}_t\sigma_v}{\sigma_v}\right)^2$, so that δ is a weighted average of the δ_t's.

Assemble these results in the full formula for σ^2_r, separating the terms involding δ, d_u, d_v, ϵ_u, ϵ_v from the others.

$$n\sigma^2_r = r^2\left\{\frac{1+2r^2}{r^2} + \frac{3}{4} + \frac{3}{4} - 3 - 3 + \frac{1+2r^2}{2}\right\}$$
$$+ mr^2\left\{\frac{\delta\sigma^2_u\sigma^2_v}{r^2\sigma^2_x\sigma^2_y} + \frac{\epsilon_u\sigma^4_u}{4\sigma^4_x} + \frac{\epsilon_v\sigma^4_v}{4\sigma^4_y} - \frac{d_u\sigma^3_u\sigma_v}{r\sigma^3_x\sigma_y} - \frac{d_v\sigma_u\sigma^3_v}{r\sigma_x\sigma^3_y} + \frac{\delta\sigma^2_u\sigma^2_v}{2\sigma^2_x\sigma^2_y}\right\}.$$

Remembering that $m\sigma^2_u = \sigma^2_x$, and $m\sigma^2_v = \sigma^2_y$, we have

$$n\sigma^2_r = (1-r^2)^2 + \frac{1}{m}\left\{\delta\left(1+\frac{r^2}{2}\right) + (\epsilon_u+\epsilon_v)\frac{r^2}{4} - (d_u+d_v)r\right\}.$$

Hence, if m (the number of independent elements) is considerable, or if the frequency curves and surfaces of the u's, v's, and pairs u_t, v_t approach normality, the correction to the usual formula $\sigma_r = \frac{1-r^2}{\sqrt{N}}$ is slight, and in any case is of the order $\frac{1}{2m}$.

If m is small, the correction may be considerable, and there is nothing to show whether it is positive or negative in general.

Some further light is thrown on the question by taking the simplified case where p out of the m u's are coincident with p out of the m v's; thus

$$x = u_1 + u_2 + \ldots + u_p + U_{p+1} + \ldots + U_m$$
$$y = u_1 + u_2 + \ldots + u_p + w_{p+1} + \ldots + w_m.$$

Here all the u's, U's, and w's are independent of each other.

Simplify further by assuming that all the u's, U's, and w's have identical frequency curves, with standard deviation σ, and that $|u^4| = |U^4| = |w^4| = (\beta_2 - 3)\sigma^4$, where β_2 is Professor Karl Pearson's measurement of "kurtosis," and equals 3 if the curves are normal. $\beta_2 - 3$ is equivalent to ϵ above.

In this case $\sigma^2_x = \sigma^2_y = m\sigma^2$ and $r\sigma_x\sigma_y = p\sigma^2_u$, so that $r = \frac{p}{m}$.

It will be found that $\rho_1 = \ldots = \rho_p = 1$, $\rho_{p+1} = \ldots = \rho_m = 0$, ${}_td_u = {}_td_v = \delta_t = \beta_2 - 3$ for p values of t and are zero for $m-p$ values, and if the simplified working is carried out the result is

$$n\sigma^2_r = (1-r^2)^2 + \frac{\kappa_2 - 3}{2m} r(1-r)(2-r).$$

For example if x is formed by the sum of two digits taken at random (0, 1, . . . 9 being equally probable) and y is formed by adding the first digit included in x to another taken at random, $m=2$, $r=\frac{1}{2}$, $\kappa_2=1.776$, say 1.8.

$$n\sigma^2_r=\frac{9}{16}-\frac{3}{32}\times 1.2=\frac{9}{16}(1-.2)=.45,$$

and $\sigma_r=\frac{.67}{\sqrt{n}}$ instead of $\frac{.75}{\sqrt{n}}$ as it would be by the usual approximate formula.

It may be added that if m and n are considerable, and x and y are formed as in the general case above, and r is not near 0 or 1, the normal table of chances is applicable to determine the probability that the deviation of r found from a sample does not exceed any assigned amount.

AN EXPRESSION FOR THE ASYMMETRICAL TENDENCY OF FREQUENCY DISTRIBUTIONS

By Stephen Varga, *Director, Hungarian Institute for Economic Research, Budapest*

Among the many existing definitions of statistics there is one maintaining that it is the science, and another declaring that it is the method of averages. These definitions are on the one hand just as serviceable and on the other about as loosely fitting as most of the others, though the point emphasized by them should be given more consideration and care than it commonly receives. There seems to be a trend—especially in the United States—to develop statistical method into an independent science,[1] a sort of applied logic or mathematics, with the primary task of determining connections and sequences of causes and effects in different fields of phenomena, on a mere statistical basis. That is, the tendency is to apply mechanically the peculiar statistical methods, without very much reference to the theories and assumptions of the different sciences, whose field of research they are invading. This makes statistics something much less inductive and much more deductive than the statisticians commonly will admit. The empirical proof for this assertion is the growing number of "statistical" studies in various fields, in contrast to the diminishing number of research works in the different sciences using statistical methods.

It is unusual to use anything but averages and ratios in descriptive economics and in economic geography, though these measures give only an inadequate notion of the central tendency [2] in the observed statistical data. To overcome the deadpoint of the average, measures of dispersion and of the symmetry or asymmetry of the frequency distribution should be employed in addition. In a recent article I emphasized the importance of the determination of the standard deviation for this purpose [3] and in a subsequent paper I used the standard deviation and the coefficient of variability to present statistically certain economic conditions. I made an investigation of the influence of the post-war agrarian reform in Hungary upon the distribution of real property between large and small owners. The result was startling inasmuch as it showed that while the share of small holdings in the entire surface

[1] I am using the expression as a slogan, with reference to the book of Hugo Forcher, *Die Statistische Methode als selbständige Wissenschaft*, Leipzig, 1913.

[2] It does not seem justified to limit the content of this expression to averages, as is done, for example by Truman L. Kelley *Statistical Method*, New York, Macmillan, 1924.

[3] Stefan Varga, *Uber die statistische Darstellung von wirtschaftlichen Verhaltnissen und Veränderungen*, Jahrbücher für Nationalökonomie und Statistik, 124 Band, III. Folge 69. Band, 1926. pp. 591 ff.

increased everywhere, the standard deviation became larger in some parts of the country. The probable reason for this development revealed itself in the different kinds of land, some of which cannot be divided into small holdings. Especially is this true in connection with forest lands.[1] It is true that it would have been possible to demonstrate this highly important *increase in the relative inequality of the size-distribution* of the real property by other methods, but it is unlikely that it would have been possible to express this result with the same certainty, exactitude and brevity—at least for those who have some notion of and feeling for mathematics—because there will always exist people and also good economists to whom mathematical and statistical formulae will not mean anything.

The importance of σ for the described purpose seems evident. But it has to be considered that σ itself, though not a strict mathematical average, yet resembles an average. This is never emphasized. It is true that some measures of skewness make use of σ in their computation, but that takes place just for the purpose of obtaining a ratio, expressing the relative importance of the asymmetry of the statistical series with regard to σ. Moreover, all measures of skewness have some defect. Some are very complicated, others can be computed only if, in addition to the mean, some other central value is also known, whereby it may occur frequently that it will be rather difficult to determine the latter.

Leaving aside Pearson's measure of skewness,[2] which is much too complicated to come into general application, the measures of skewness, the use of which is conventionally suggested, are the following:

$$\text{Skewness} = \frac{\text{arithmetic mean} - \text{mode}}{\text{standard deviation}}$$

$$\text{Skewness} = \frac{\text{upper quartile} + \text{lower quartile} - 2\ \text{median}}{\text{upper quartile} - \text{lower quartile}}$$

and $\text{Skewness} = \text{median} - \frac{1}{2}\ (\text{percentile}_{90} + \text{percentile}_{10})$.[3]

It must be emphasized that these formulae are scarcely ever applied outside of theoretical introductions into the methods of statistics, with the possible exception of biometry. There is good reason for this. The frequency distribution in economic statistical data, for example, is often irregular. It is practically certain that in cases of bimodal

[1] "The Agrarian Reform and the Inequality in the Distribution of Landed Property," *Hungarian Economic Review* (In Hungarian), Budapest, 1926, pp. 25 ff.

[2] See *Handbook of Mathematical Statistics*, edited by H. L. Rietz, Houghton Mifflin Company, 1924, Chap. VII on "Frequency Curves," by H. C. Carver, pp. 103–104.

[3] For some other measures of skewness see G. Udny Yule, *An Introduction to the Theory of Statistics*, London 1911, pp. 149–150.

frequency distributions, the mode can scarcely be used, *i. e.*, in most cases its determination will be intractable. Also there will be difficulties with the mode when we have frequency distributions of a type in which the frequencies in the classes of the statistical series increase or decrease continuously, as quite often in economic statistics. The mode will be in these cases within the first or last class, and will not be influenced by the form of the frequency distribution in the other parts of the array. It will happen also rather often that the class intervals of the published statistical tables—in cases when secondary sources are used—are rather large and the number of classes, consequently, small. The cognitive value of the mode will be small in the latter case and the defect can scarcely be remedied through interpolation. Some of the imperfections will influence the determination of the median and of the other percentiles as well. It is true that interpolation may be used to secure closer approximation for these. But this will mean additional work and it will always remain true that neither median nor mode nor any of the percentiles is influenced by the entire frequency distribution. That is, in making use of them, we may get the same result, the same number, as an expression of the skewness of two or more differently asymmetrical statistical series. And this possibility diminishes the usefulness of these measures considerably.

To get a more useful and at the same time more easily computable measure to express the tendency of asymmetry in statistical series it is justifiable to start with the previously emphasized understanding that the standard deviation, which has always to be determined, is itself something like an average. To get an expression for the tendency of asymmetry we can break its computation into two parts. The standard deviation is, as everybody knows, $\sigma=\sqrt{\frac{\Sigma X^2}{N}}$, where X stands for the deviations from the mean, and N is the number of units in the series. If we take the minus and plus deviations separately, we can say that the quoted familiar formula of

$$\sqrt{\frac{\Sigma X^2}{N}}=\sqrt{\frac{\Sigma(-X)^2+\Sigma(+X)^2}{N_1+N_2}} \tag{1}$$

where N_1 equals the number of data smaller and N_2 the number of data larger than the mean. Now it is obvious that σ is not the average of

$$\sqrt{\frac{\Sigma(-X)^2}{N_1}} \text{ and of } \sqrt{\frac{\Sigma(+X)^2}{N_2}}, \text{ as}$$

$$\sqrt{\frac{\Sigma(-X)^2+\Sigma(+X)^2}{N_1+N_2}} \neq \frac{\sqrt{\frac{\Sigma(-X)^2}{N_1}}+\sqrt{\frac{\Sigma(+X)^2}{N_2}}}{2}. \qquad (2)$$

But taking the mean as starting point for the determination of the deviations, the difference will not be very great in most cases (if there is perfect symmetry, the inequality (2) becomes naturally a perfect equation). And it is obvious at the same time that though no average of $\sqrt{\frac{\Sigma(-X)^2}{N_1}}$ and of $\sqrt{\frac{\Sigma(+X)^2}{N_2}}$, σ is influenced by their ratio. This justifies our taking for a sound expression of the asymmetrical tendency of the frequency distribution of a statistical series the ratio

$$\sqrt{\frac{\Sigma(-X)^2}{N_1}} : \sqrt{\frac{\Sigma(+X)^2}{N_2}}$$

or, designating these values by γ_1 and γ_2 respectively the ratio of $\gamma_1:\gamma_2$ The ratio will become still more explicit if we express it on a per cen basis, in such a way that $\epsilon_1+\epsilon_2$ shall equal 100, where $\epsilon_1:\epsilon_2=\gamma_1:\gamma_2$. We may say then *that σ is influenced through the asymmetry of the frequency distribution in the statistical series, or, what amounts to the same, through the negative and positive deviations of the data from the mean in the ratio of $\epsilon_1:\epsilon_2$ in the negative and positive direction, respectively.* In case of perfect symmetry $\epsilon_1:\epsilon_2$ is equal to 50:50.

Two cases must be taken into consideration relative to the computation of γ_1 and γ_2, or of ϵ_1 and ϵ_2. It may happen that the deviation of some class from the mean is zero. In this case it is advisable to take into account the relative class frequency in the computation of both γ_1 and γ_2. It will frequently happen also that the standard deviation is computed for simplicity—to avoid too many fractions, for example—starting with an arbitrary origin. This presents a difficulty. Yet, if this arbitrary origin is chosen in a way to be close to the mean—and that will always be possible—it will influence the proportion of γ_1 and γ_2 very little, so that it does not seem necessary to make any correction.

The proposed expression for the tendency of asymmetry in a frequency distribution (for it seems better to avoid speaking of a *measure*, which would imply more accuracy for it than it justly can claim for itself) is certainly not perfect, and its mathematical imperfections have been pointed out. But on the other hand, it has the merit of simplicity. And if in some cases it distorts the asymmetry, making it seem greater than it is, that will not greatly influence its usefulness for comparisons.

Let us not forget that many other statistical measures of central tendencies also lack mathematical exactitude in practical use.

To compute γ_1 and γ_2 in addition to σ does not present any difficulty and can always be done with very little additional work. And to repeat: The only way to overcome the deadpoint of mathematical abstraction in the simplified expressions of central tendencies in frequency distributions of statistical series is to add an expression for their asymmetry to the determination of their means and their standard deviations.

A STATISTICAL STUDY OF THE ECONOMIC AND SOCIOLOGICAL SIGNIFICANCE OF POPULATION AGE[1]

By Nathan P. Levin, *Los Angeles County Health Department*

It is a fact that current economic doctrine is largely preoccupied with such material things as money, railroads, crops, and wealth in general. But before any other component of national wealth we should consider the composition of the population itself. It is the most important because it is the most fundamental of wealth-producing factors. I am not referring to mere numerical magnitude of a population. This, of course, enters as a common factor in nearly all economic calculations. Rather do I wish to emphasize the importance of the age group distribution within the population. Two populations equal in numbers, equal in wealth, equal in every respect save age group distribution, will not be economically equivalent. In what follows I propose to show how it is possible to measure the economic value of any age group distribution, and statistically represent it by a definite number. This particular relationship between population and economics seems to have been quite generally overlooked.

In all trades and occupations such as farming, bricklaying, machine tending, etc., where physical effort is required, earning capacity is a function of age. Even in the small group of occupations, predominantly mental, age is a factor not to be ignored in estimating probable earning power. Plainly, it should be possible by statistical methods, to derive index numbers representing average relative earning capacities for each age. Despite the great theoretical and practical value of such a formulation, it has been attempted in only one instance, namely by a committee of the International Association of Industrial Accident Boards. Its primary aim was to derive a better basis for total disability compensation. The figures are reproduced in Table I. In Table II these figures have been averaged by five-year age intervals, to correspond with composition of population statistics as ordinarily given in census tables. In answer to my inquiry regarding the basis of computation of this table, Mr. Ethelbert Stewart, United States Commissioner of Labor Statistics, replied as follows:

> . . . this table of index numbers was pioneer work, as there were no statistics covering actual experience, on which to base scientific tabulations. It was felt, though, that there was a direct need for something that could be used as a

[1] *Based in part upon thesis submitted to the University of California for the degree of Dr. P. H.*

TABLE I

INDEX NUMBERS SHOWING RELATIVE ANNUAL EARNING CAPACITY OF WORKMEN BY ORDER OF AGE

The age of maximum earning, 40, is taken as the standard, 100
From the *United States Bureau of Labor Statistics Bulletin No. 359*

Age	Index of earning	Age	Index of earning
15	50	45	95
16	53	46	94
17	56	47	93
18	59	48	92
19	62	49	91
20	65	50	90
21	68	51	88
22	71	52	86
23	74	53	84
24	77	54	82
25	80	55	80
26	82	56	78
27	84	57	76
28	86	58	74
29	88	59	72
30	90	60	70
31	91	61	68
32	92	62	66
33	93	63	64
34	94	64	62
35	95	65	60
36	96	66	58
37	97	67	56
38	98	68	54
39	99	69	52
40	100	70	50
41	99	71	48
42	98	72	46
43	97	73	44
44	96	74	42

TABLE II

AVERAGE ANNUAL EARNING CAPACITY FOR FIVE-YEAR AGE INTERVALS

Computed from Table I

Age interval	Average earning capacity *
15–19	.56
20–24	.71
25–29	.84
30–34	.92
35–39	.97
40–44	.98
45–49	.93
50–54	.86
55–59	.76
60–64	.66
65–69	.56
70–74	.46

* The index numbers are given as decimals, instead of whole numbers of the original figures. This is because, for the purposes of our study, it is convenient to posit unit earning capacity equal to one instead of one hundred.

guide in the determination of equitable compensation to injured persons of different ages, therefore this table of index numbers was prepared without any claims of scientific accuracy, but merely as a guide to those administering workmen's

compensation regulations. This table represents the best judgment of the Committee on Statistics and Compensation Insurance Cost, of the International Association of Industrial Accidents Boards and Commissions, which was composed of the foremost workmen's compensation statisticians of the United States. This table is the only attempt that I know of that has ever been made regarding the matter.

It follows from Table II that the productive capacity of a population depends not only upon numbers, but also upon age distribution within the population. It is desirable to be able to compare different populations as regards economic potentialities due to the one factor of age composition. For this purpose I shall make use of a magnitude to be known as "basal productivity" defined as the sum of the earning indices of every individual in the population, divided by the total of the population.

If P numerically represents the total population, and A the number of individuals in a definite age group of that population, and E the index number of earning capacity corresponding to the age group, then

$$\text{basal productivity} = \frac{\Sigma(AE)}{P}.$$

Basal productivity as thus defined is, therefore, a true statistical mag-

TABLE III

COMPARATIVE BASAL PRODUCTIVITY AND COMPARATIVE BASAL EFFICIENCY AS AFFECTED BY AGE, UNITED STATES POPULATION, BOTH SEXES, 1910

a Age	b Persons per million	c Earning index	b×c Earning capacity	d Food index	b×d Food requirement
0– 4	115,800	0	0	.37	43,400
5– 9	106,300	0	0	.48	51,024
10–14	99,200	0	0	.60	59,520
15–19	98,700	.56	55,272	.69	68,103
20–24	98,600	.71	70,006	.86	84,796
25–29	89,100	.84	74,844	.90	80,190
30–34	75,900	.92	69,828	.90	68,310
35–39	69,700	.97	67,609	.90	62,730
40–44	57,300	.98	56,154	.90	51,570
45–49	48,700	.93	45,291	.90	43,830
50–54	42,500	.86	36,550	.90	38,250
55–59	30,400	.76	23,104	.90	27,360
60–64	24,700	.66	16,302	.75	18,525
65–69	18,300	.56	10,248	.75	13,725
70–74	12,100	.46	5,566	.75	9,075
75–79	7,300	.36	2,628	.75	5,475
80–84	3,500	.26	910	.75	2,625
85–89	1,300	.16	208	.75	975
90–94	400	.6	24	.75	300
	1,000,000		534,544		729,783

Comparative basal productivity = average earning capacity per person
$= \Sigma(b \times c)/\Sigma b = 53.4$

Comparative basal food requirement = average food consumption per person
$= \Sigma(b \times d)/\Sigma b = 73.0$

Comparative basal efficiency = basal productivity divided by basal food consumption
$= \Sigma(b \times c)/\Sigma(b \times d) = .732$

nitude, representative of the entire population and changing as the age composition of the population changes.

The method of computation is exemplified in Table III, wherein are tabulated the figures for the derivation of the basal productivity of the United States population, Registration Area 1910. Column "a" tabulates the age groups by quinquennial intervals; column "b" the number of individuals in each million of population corresponding to each age interval; column "c" the average earning index for the corresponding age interval, the index corresponding to age 40 being taken as unity; while column "b×c" is self-explanatory.

By this method I have computed the basal productivities of some five populations, as follows:

Germany	1900	50.9
United States Registration Area	1910	53.4
*United States Stationary	1910	58.3
France	1900	57.3
California	1920	61.4

For the United States basal productivities have been computed both for actual population and stationary population, for the year 1910. The stationary population, it will be remembered, is computed to show no increase or decrease on the basis of actual mortality rates prevailing at the time. The figures show, as we would expect, a much higher basal productivity for the stationary population. It is 58.3 as compared with 53.4. It is therefore 9 per cent higher. A growing population must always show a smaller basal productivity than a stationary one. The provision for growth requires a proportionately greater number of individuals in the younger age groups, and these reduce the basal productivity because of their lower earning power. We should look upon this 9 per cent as the proportion of potential earnings of the nation, sacrificed to provide for growth, over and above mere maintenance.

It will be very instructive to devote some space to a consideration of the significance of the basal productivities of the two European nations in the list. Important social and political consequences as well as economic ones will be deduced. The case of France is an example of a population which is actually stationary. From 1860 to 1913 the birth-rate of France has been the lowest in Europe. This has brought about

* The "stationary population" United States Registration Area 1910, is computed in Table 2, *United States Life Tables*, Washington, 1921. It is there defined as "unaffected by emigration and immigration, which assuming the mortality rates in column 4, would result if 100,000 persons were born alive uniformly throughout each year". Under these special conditions the population would eventually attain a constant or stationary level as to numbers; and what is more important from our point of view, its composition as to age of individuals would likewise remain constant. "Stationary population" is therefore, not an actual, but computed ideal population.

a preponderance of persons of higher age. As long ago as 1876 the average age of all living in France was 32.6 years, compared with 29.2 for Switzerland, and 26.0 for England.

In consequence of its stationary character the French population shows a high basal productivity, namely 57.3. This is very close to that of the United States stationary population with a magnitude of 58.3. It is instructive to look upon the population problem in France as a vast experiment in the national conservation or economizing of human capital. We have become accustomed to think only of material things as capital. But any human being, at any time of his life, is very definitely either an economic asset or liability to the nation, and in so far as this is measurable in terms of dollars and cents, human beings are to be regarded as capital. The reduction of the birth-rate in France may be an attempt at the reduction of national overhead in developing human material.

In 1911 in France there were 11,700,000 families averaging 3.3 persons per family. The average of dependents under 15 years of age was .95 per family. On the other hand, in Germany in 1905 there were 4.7 persons per family, 2.29 of which were dependents under 15 years of age. In other words, the average French family was saddled with one and one-third less dependent children than the average family across the Rhine. Now certain studies which I have made [1] lead to the conclusion that in the United States the minimum cost of child upkeep averages about $200 annually, while a study made by the statistical department of the Metropolitan Life Insurance Company [2] leads to a figure more than double this amount. But even on the basis of the lesser figure a simple calculation will show that, translated in terms of United States currency, and with post-war price levels, three billions of dollars is the least amount saved annually by the French people by mere virtue of their population being stationary instead of growing at the same rate as that of Germany.

In the light of these facts we are no longer surprised to learn that the per capita wealth of France before the War was the highest in Europe, and that her total national income estimated by Mulhall for 1895 [3] was only 6.6 per cent less than that of Germany, notwithstanding a 45.0 per cent difference in population.

The highest basal productivity in our list is 61.4, applying to the 1920 population of the state of California. This is in keeping with the composition of its population which is even older in type than that of

[1] N. P. Levin, Thesis, *A Critical Study of the Economic Background to Public Health*, University of California, Berkeley.

[2] Louis Dublin, "Economics of World Health," *Harpers Monthly*, November, 1926.

[3] *Abstract of the Fourteenth Census*, p. 186.

pre-war France. Only 38.8 per cent of the California population is under 25 years of age, as compared with the 38.8 per cent of France and 50.0 per cent for the United States as a whole. To be sure, the influences responsible for this type of population are quite different in the two countries. In the case of France it is brought about by means of voluntarily controlled fertility, whereas in the case of California the diminution in the younger ages is relative, and the consequence of the tremendous accretion in the more mature age groups brought about by immigration. Of the California population of 1920, 62.7 per cent were born outside of the limits of the state. The high basal productivity of California, the consequence of this immigration, is one of the most important sources of its great wealth.

California exemplifies a type of immigration which may be termed "internal," that is from one part of the same country to another. From the economic standpoint, the most important type of internal migration is not from state to state, but from country to city. According to the United States census for 1920, the percentage of urban population less than 20 years of age, is 35.8, while in rural United States this group constitutes as much as 45.9 per cent of the total.[1] It follows that the basal productivity of the urban population is much higher than that of the rural. The difference in basal productivity between city and country probably reflects the fact that it is economical to raise human livestock on country pastures, up to the age when the individuals become useful for energy production in the city factories.

In the foregoing discussion of the economics of population, I have considered at some length the productive capacity of population groups as conditioned by their age composition. I have shown how this characteristic for any population of known composition can be expressed in terms of a single number known as basal productivity. But this is only half of the picture. In any study involving economics we are interested not only in production but also in consumption. To make my treatment complete, I must supplement the foregoing with a parallel study of consumption requirements of a population as conditioned by its age composition.

For an index of consuming force we are practically limited in choice to the one item, food. We may feel sure that it is adequate for this purpose. First, because of all human needs it is the most indispensable; second, the food requirement at different ages lends itself to quantitative evaluation; and third, because the expenditure for food constitutes a large and fairly constant proportion of total expenditure for

[1] *Ibid.*, p. 186.

necessities. For the bulk of the productive population of the United States about 43 per cent of the family budget goes for food.[1]

There are two sets of index numbers for relative amounts of food consumed at different ages.

TABLE IV

INDEX NUMBERS OF RELATIVE FOOD REQUIREMENTS AT DIFFERENT AGES

Age	United States Bureau of Labor statistics	Atwater *
1	.15	.3
2	.15	.4
3	.15	.4
4	.4	.4
5	.4	.4
6	.4	.5
7	.75	.5
8	.75	.5
9	.75	.5
10	.75	.6
11	.90	.6
12	.90	.6
13	.90	.75
14	.90	.75
15	.95	.75
16	.95	.85
17 and over	.95	.90

* W. O. Atwater in United States Department of Agriculture *Farmers Bulletin No. 142*. In the above table I have averaged his figures for males and females.

The unit for the United States Bureau of Labor Statistics indices is the average food consumption of individuals 15 years of age, while the unit of the Atwater figures is an able bodied man at moderate work. The figures in Table IV are all less than unity because they represent the average of both male and female indices. In the original figures the indices are given separately for the two sexes, and the figures for the females run about .10 units less than those for the males.

The two scales differ somewhat because they are computed from entirely different data and points of view. The United States Bureau of Labor Statistics numbers are derived from family budgets, while Atwater's figures represent food requirements based on physiologic consideration of metabolism studies. After full productive age is attained the food requirement remains constant until considerably past middle age. With elderly individuals, however, "the intensity of metabolism is diminished so that food requirement gradually declines, and may become 10 per cent, or 20 per cent or 30 per cent lower than in middle life."[2] I have selected the Atwater scale for use in this population economics study, as it is strictly scientific and based on invariable biologic factors, whereas the United States Bureau of Labor

[1] *United States Bureau of Labor Statistics, Bulletin No. 357.*

[2] H. C. Sherman, *Chemistry of Food and Nutrition*, 2d edition, p. 193.

Statistics series conceivably may differ with changes in relative food prices from time to time, and place to place. To conform with the observation of Sherman just quoted, I have assumed on this scale an average index number of .75 for all ages over 60. From the Atwater scale thus supplemented, I have computed relative food requirements by quinquennial age periods. This is listed in Table III, column "d".

It is now necessary to define a magnitude analogous to basal productivity, which will have the same relationship to the consuming power of a population that basal productivity has to the productive power. This magnitude is called "comparative basal food requirement." It is defined as the sum of the food consumption indices of every individual, divided by the total of the population.

If P represents numerically the total population, and A the number of individuals in some definite age group of the population, and F the average relative food consumption for that age group, then

$$\text{basal food requirement} = \frac{\Sigma(AF)}{P}.$$

For the example we have chosen in Table III, basal food requirement comes to 73.0.

What we are mainly interested in is the ratio of human resources to human needs, or of production to consumption. In line with this a magnitude to be known as "comparative basal efficiency" is defined as simply the ratio of basal productivity to basal food consumption. By itself this number possesses no significance whatsoever. Its significance is entirely relative or comparative. Of two populations the one with the higher basal efficiency is in a material sense the better situated, because it possesses a greater reserve of human resources of production. After the satisfaction of elemental needs, there will be a greater balance in favor of the more "basally efficient" nation, to apply towards the acquisition of savings, comforts and luxuries. It will be of interest to compare the populations previously listed, as regards their basal efficiencies.

COMPARATIVE BASAL EFFICIENCIES

Germany	1900	.707
United States Registration Area	1910	.732
United States Stationary	1910	.770
France	1900	.763
California	1920	.800

It is at once apparent that for the group of populations studied, the variability in basal efficiencies is less than that in basal productivities. This must be true in general, owing to the fact that food requirement

tends to keep pace with productivity or earning capacity; both increasing with age up to a certain point. Consequently their quotient tends to fluctuate around a constant.

In this paper has been indicated a new method for the quantitative study of certain important economic and social problems connected with population. I do not wish to emphasize numerical results, which ought and must suffer modification as soon as more reliable and refined basic data are available.

AN AGGREGATIVE INDEX OF FARM PURCHASING POWER

By Forrest Bee Ashby, *University of Pennsylvania*

Periods of agricultural depression are prolific in economic discussions of farm problems. Thus the post-war difficulties of American farmers have given rise to numerous books and magazine articles dealing with the troubles of this class of our citizens.[1] The majority of these writers arrive at the conclusion that the farmer's economic situation is deplorable, although there are some exceptions, such as Mr. David Friday.

What concerns us here, however, is the statistical basis for these conclusions. It is, in almost all cases, a comparison of the trend of agricultural income with the trend of agricultural outgo, as shown by a comparison of the price indexes of farm products with the price indexes either for all commodities or for certain non-agricultural commodities. The all-commodities price index is usually the Bureau of Labor figure and as such is weighted according to the value of the 400-odd articles dealt in during 1909. The price index of farm products is sometimes an unweighted arithmetic average of selected agricultural prices (such as wheat, corn and cotton), but is often weighted by the relative importance (according to value) of the constituent series, the quantities being chosen as of any year, but, for some reason, usually 1919.[2]

It is desired here to make only one point, and that is, as has been frequently admitted, that the quantity of his crops determines the farmer's income or purchasing power just as much as the price of each crop unit. This is to say, in other words, that an index of farm income or farm purchasing power should be of an aggregative type, where the price of each period for each commodity is multiplied by the marketed quantity of that commodity, and the products totalled. Yet such an index is rarely used,[3] one writer even regretting the fact that data for its construction are not available. It is the purpose of this paper to

[1] Outstanding are *The Agricultural Situation*, by G. F. Warren and F. A. Pearson; "The Effect of Price Fluctuations on Agriculture," A. H. Hansen (*Journal of Political Economy*, Vol. 33, 1925, p. 196); "The Course of Agricultural Income," David Friday (*American Economic Review*, Vol. 13—supplement, 1923, p. 147); "The Agricultural Depression," G. F. Warren (*Quarterly Journal of Economics*, Vol. 38, 1924, page 183); "The Problem of Agriculture," R. G. Tugwell (*Political Science Quarterly*, Vol. 39, 1924, p. 549).

[2] See *Review of Economic Statistics*, 1922, p. 224; *Quarterly Journal of Economics*, Vol. 38, p. 188, "The Agricultural Depression," by G. F. Warren.

[3] Only once to the writer's knowledge. See "Agricultural Income," David Friday, *American Economic Review*, Vol. 13—supplement, p. 153. And Mr. Friday uses only figures for annual gross value of United States farm products.

illustrate such a barometer, to give the sources of the data used in its compilation, and to suggest uses therefor.

The compilation of the aggregative index of farm purchasing power is as follows: In the first place, the constituent series must be chosen. The writer chose seven: Wheat, corn, oats, cotton, cattle and calves, hogs, and sheep. These represent about two-thirds of the marketed farm produce in America, and omit only two important commodities, milk and hay. A commercial index should include these two items. The sheep are negligible in importance and may be omitted. In the second place, the price for each product was listed by months over the period 1913–1926. In the third place, each monthly price was multiplied by the quantity of that product entering the American market.[1] In the fourth place, the aggregate values of all the commodities were totalled month by month. The monthly figures were used to obtain quarterly figures, which are sufficiently flexible to indicate economic fluctuations, but are less likely than monthly figures to vary erratically.

These figures, which may be called prime data, may be calculated period by period, and may be manipulated to suit the needs of the economic investigator or business statistician. They are shown in Chart I, reduced to relatives of a 1913 average, and against these figures appears the Bureau of Labor revised price index of farm products, which is also on a 1913 base.

[1] Sources of data were:

Wheat: Receipts—1913–June, 1921, *Market and Crop Reporter*, United States Department of Agriculture.
July, 1921–October, 1921, the monthly receipts at the primary markets, equated to the (later) receipts at points of inspection. This period was one of change in the method of reporting receipts.
November, 1921–1926, monthly receipts in carload lots at all points of inspection; estimated carload capacity, 1500 bushels. Reported by *Market and Crop Reporter*.
Price— The United States average price paid to producers per bushel on the 15th of each month. *Market and Crop Reporter*.

Corn: Receipts—Same. Estimated weight of carload, 1800 bushels.
Price— Same.

Oats: Receipts—Same. Carload lots were weighed to bushels of the estimated carload capacity, 2,000 bushels.
Price— Same.

Cotton: Receipts—Bales into sight as reported from 1913 to 1924 by the *Journal of the New Orleans Cotton Exchange*, and 1925–26 by the *Survey of Current Business*, United States Department of Commerce.
Price— Average monthly prices paid to producers per pound, as reported by the *Market and Crop Reporter*. Estimated: 500 pounds per bale.

Cattle: Receipts—1913–1918, yearly receipts at public stockyards of United States, reduced to monthly figures on basis of percentage of marketings monthly at the Chicago market in 1911. Source, *Year Book, Chicago Board of Trade*.
1919–1926, monthly receipts at public stockyards, taken from *Market and Crop Reporter*.
Price— Average monthly, *Market and Crop Reporter*. Estimated weight, 975 pounds; stockers and feeders included.

Sheep: Source of receipts and price same as for cattle. Weight 85 pounds.

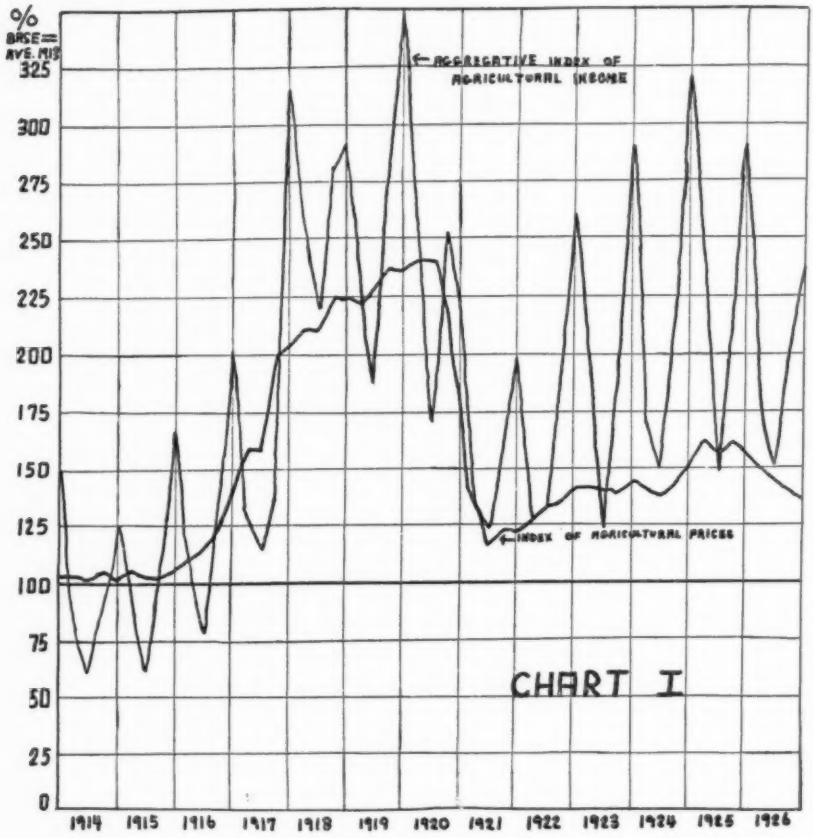
Hogs: Source of all information, *Market and Crop Reporter*. Weight estimated at 220 pounds.

AN AGGREGATIVE INDEX OF FARM PURCHASING POWER

Period	Values paid to farmers (000 omitted)	Relatives of payments *	Wholesale price index of farm products *
1913			
First quarter	$612,990	87.8	98.0
Second	445,811	64.0	98.7
Third	677,471	97.1	100.7
Fourth	1,043,887	150.0	103.0
1914			
First quarter	627,917	90.0	102.7
Second	416,525	59.8	101.3
Third	675,501	97.0	105.0
Fourth	870,714	125.1	101.3
1915			
First quarter	703,266	101.0	105.0
Second	464,514	66.8	103.3
Third	750,861	108.0	102.7
Fourth	1,162,416	167.6	105.3
1916			
First quarter	781,104	112.2	110.3
Second	560,606	80.7	114.3
Third	939,791	135.2	124.3
Fourth	1,436,828	203.4	143.0
1917			
First quarter	933,599	134.2	158.3
Second	801,042	115.1	158.3
Third	950,736	137.0	200.0
Fourth	2,208,261	317.0	204.7
1918			
First quarter	1,788,582	257.0	211.3
Second	1,525,496	219.2	210.7
Third	1,962,362	282.2	226.0
Fourth	2,036,162	293.0	225.0
1919			
First quarter	1,525,485	219.2	221.3
Second	1,290,457	185	230.0
Third	2,018,996	290.8	236.0
Fourth	2,414,098	347.0	235.3
1920			
First quarter	1,676,832	241.3	240.3
Second	1,176,698	169.2	240.3
Third	1,764,505	254.0	220.3
Fourth	1,483,249	213.3	170.7
1921			
First quarter	925,788	133.2	134.3
Second	867,311	124.8	116.3
Third	1,179,026	169.8	122.0
Fourth	1,381,825	199.0	121.7
1922			
First quarter	874,700	125.8	127.7
Second	952,625	137.0	130.7
Third	1,343,424	193.2	133.0
Fourth	1,815,636	261.0	142.0
1923			
First quarter	1,253,940	180.3	142.7
Second	854,119	122.8	139.3
Third	1,380,618	198.7	139.3
Fourth	1,974,946	284.0	145.4
1924			
First quarter	1,188,582	171.0	141.3
Second	1,043,832	150.3	136.3
Third	1,536,877	221.0	143.0
Fourth	2,211,819	318.0	152.0
1925			
First quarter	1,446,626	208.0	162.0
Second	1,029,884	148.0	153.0
Third	1,608,243	231.5	161.7
Fourth	2,037,429	293.9	153.7
1926			
First quarter	1,231,909	177.2	149.0
Second	1,043,327	150.2	144.0
Third	1,425,916	205.0	140.0
Fourth	1,656,228	238.1	136.0

* Base for relatives average of 1913.

It will be noticed that there is a tremendous seasonal element in the aggregative purchasing power index in its unadjusted condition, while the seasonal in the index of agricultural prices is not apparent. The graph shows that up to and including 1920 the agricultural-commodity price index represented rather fairly the general swings of the agricultural purchasing power index. After 1920, however, the agricultural-commodity price index consistently falls far below the aggregative



purchasing power index. This means that in the later years a greater quantity of crops was being marketed.

It may be said that this increase in marketings is due to a greater number of persons engaged in a more intensive cultivation of the soil. Few figures are available on this subject. We know that in 1920 there were some six million farmers engaged in agriculture, and that they were assisted by four million farm hands. In 1922 the Secretary of Agriculture estimated that, in that year alone, 1,200,000 persons left

the farms for the cities, which, together with the prevailing pessimism regarding agriculture as an occupation, would argue strongly that the number of persons engaged in agriculture has increased little if at all since 1920. This would lead to the conclusion that the per capita purchasing power of farmers since 1920 has been far above that indicated by an inspection of mere agricultural-commodity price indexes, and that this higher income is due to the greater quantities of product marketed.

The influence of the aggregate of dollar purchasing power must be modified by other factors in studying such problems as, for instance, the current and expected market conditions for manufactured articles sold nationally to farmers. Some of these other factors are changes in rural population (though changes in this item would seem to be slow at the present time), changes in farm taxes, and changes in the burden of interest owed on farm obligations, and also changes in the price level of manufactured goods themselves. As between two periods of time, assuming farm overhead and farm population unchanged, the trend of farm demand for non-agricultural commodities should be indicated by the trend of the quotients resulting from the calculation for each period of the fraction:

$$\frac{\text{Aggregate farm income}}{\text{Price level of non-agricultural products.}}$$

It is evident that in studying agricultural prosperity, farm purchasing power and similar problems, a number of indexes are needed which are not now available, such as an index of per capita income, both cash and real; and also indexes of the various forms of outgo, such as labor, interest and so on. It is the writer's intention here merely to suggest that such an index of cash income should be represented, period after period, by the amount of dollars received by the farm group and available for later expenditures, instead of by an index of agricultural prices alone.

As some evidence to support this conclusion, Chart II is presented, showing a graphic comparison of the cyclical swings (that is, seasonal and secular removed[1]) of the aggregative index of farm purchasing

[1] The cycles are derived by L. W. Hall's derivative integral method, using a three-year moving average of seasonals and a four-quarter moving integration. Algebraically:

Let O equal the original data.

$$\frac{O_2}{O_1}, \frac{O_3}{O_2}, \text{ etc.} \quad = A \text{ (chain index)}$$

$$A-100 \quad = B \text{ (residual in per cent)}$$

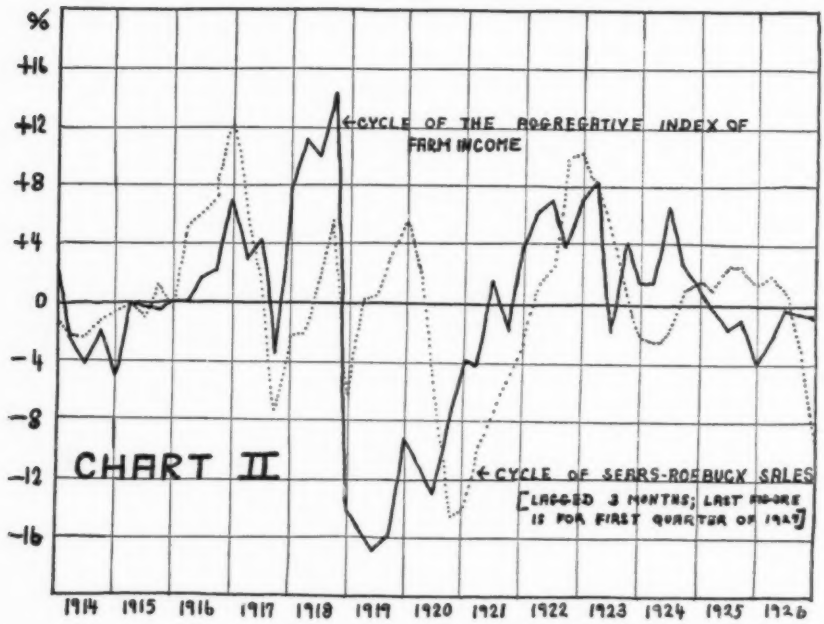
$$\frac{B_1+B_2+B_3}{3} \ldots \quad = C \text{ (three year moving average of seasonals)}$$

$$B-C \quad = D \text{ (relative derivative cycle)}$$

$$\frac{D_1+D_2+D_3+D_4}{4} \ldots \quad = E \text{ (relative derivative integral cycle)}$$

power with the sales of Sears, Roebuck and Company, national vendors of farmers' necessities, with the latter curve lagged three months.

The high correlation existing between the cyclical sales curve and the



aggregative index of farm income in its elementary form tends to indicate the value of a completely developed index of this type for agricultural purchasing power

NOTES

AN ELEMENTARY PROOF OF STIRLING'S FORMULA

By Robert E. Moritz, *University of Washington*

The belief is widespread among statisticians, and is occasionally shared by mathematicians,[1] that Stirling's Formula, viz., $n! = e^{-n}n^n\sqrt{2\pi n}$, approximately, for large values of n, is not susceptible of proof by elementary means.

For this reason, which is frankly stated by some authors, it has become the custom for writers of treatises and textbooks on statistics and statistical methods, including such eminent authorities as G. U. Yule[2] and E. Czuber,[3] to state the theorem without proof. This is to be regretted inasmuch as Stirling's Formula offers the easiest and most natural approach to the derivation of the error function which is recognized by all as the corner stone of the structure of theoretical statistics.

It is quite true that the usual proofs of Stirling's Formula require a knowledge of gamma functions,[4] or of finite differences and Bernoulli's numbers,[5] but so-called elementary proofs have been given from time to time though these proofs have not found their way into textbooks on statistical methods. This is probably because these proofs, while ostensibly elementary in character, are found, on critical examination, to involve principles which lie outside the field of elementary mathematics, if we limit this field to those branches of mathematics which precede and include the elements of the differential and integral calculus.

Thus the earliest of these proofs, given by Glaisher,[6] assumes without proof the convergency of a double series and treats as constant a variable which approaches a constant as its limit. Chrystal's[7] elementary proof makes use of one of Dirichlet's theorems on limits. Cesaro's[8] proof, reproduced by Arne Fisher in his *Mathematical Theory*

[1] *Bulletin American Mathematical Society*, Vol. 24 (1917–18), p. 478.
[2] *Introduction to the Theory of Statistics*, London (1910), p. 304.
[3] *Die Statistischen Forschungsmethoden*, Wien (1921), p. 192.
[4] Todhunter, *Treatise on the Integral Calculus*, London (1895), p. 269.
[5] Whittaker and Robinson, *Calculus of Observations*, New York (1924), p. 138; Markoff, *Differenzenrechnung*, Leipsig (1896), p. 139.
[6] *Quarterly Journal of Mathematics*, Vol. 15, pp. 57–65.
[7] *Textbook of Algebra*, London (1919), p. 368.
[8] *Corso di Analisa Algebrica*, Turino (1884), pp. 270, 480.

of Probabilities, involves certain of Weierstrass' theorems on number sequences.

It is the purpose of this paper to present a proof of Stirling's Formula which is elementary in the sense that it requires no knowledge of mathematics beyond algebra, for while, like all other proofs of the formula, it makes use of Wallis' formula for the value of π, Wallis' formula may be readily derived algebraically, though the more convenient method of definite integrals is the more common procedure.

I will state explicitly the lemmas which are necessary in the proof which follows:

(A) $\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots$, when $x \leqq 1$.

(B) $\lim_{n\to\infty} \left(1+\frac{1}{n}\right)^n = e$, the base of the Naperian system of logarithms.

(C) $(\frac{1}{2})^k + (\frac{1}{3})^k + (\frac{1}{4})^k + \ldots$, converges when $k>1$, and approaches zero as a limit when k is indefinitely increased.

(D) Wallis' Formula, $\pi = \lim_{n\to\infty} \left\{\frac{1}{n}\left(\frac{2}{1}\cdot\frac{4}{3}\cdot\frac{6}{5}\cdots\frac{2n}{2n-1}\right)^2\right\}$.

The proof follows:

$$\text{Let} \qquad u = \frac{2}{1}\left(\frac{3}{2}\right)^2\left(\frac{4}{3}\right)^3 \cdots \left(\frac{n+1}{n}\right)^n = \frac{(n+1)^n}{n!} \tag{1}$$

then by (A)

$$\log u = \log 2 + 2\log\left(1+\frac{1}{2}\right) + 3\log\left(1+\frac{1}{3}\right) + \ldots + n\log\left(1+\frac{1}{n}\right)$$

$$= \log 2 + 2\left(\frac{1}{2} - \frac{1}{2}\cdot\frac{1}{2^2} + \frac{1}{3}\cdot\frac{1}{2^3} - \ldots\right)$$

$$+ 3\left(\frac{1}{3} - \frac{1}{2}\cdot\frac{1}{3^2} + \frac{1}{3}\cdot\frac{1}{3^3} - \ldots\right)$$

$$+ 4\left(\frac{1}{4} - \frac{1}{2}\cdot\frac{1}{4^2} + \frac{1}{3}\cdot\frac{1}{4^3} - \ldots\right)$$

$$+ \ldots\ldots\ldots\ldots\ldots\ldots$$

$$+ n\left(\frac{1}{n} - \frac{1}{2}\cdot\frac{1}{n^2} + \frac{1}{3}\cdot\frac{1}{n^3} - \ldots\right)$$

$$= \log 2 + (n-1) - \frac{1}{2}\sum_2^n \frac{1}{n} + \frac{1}{3}\sum_2^n \frac{1}{n^2} - \frac{1}{4}\sum_2^n \frac{1}{n^3} + \ldots$$

$$=\log 2+n-\frac{1}{2}-\frac{1}{2}S_n+C_n, \tag{2}$$

where $S_n=1+\frac{1}{2}+\frac{1}{3}+\ \ldots\ +\frac{1}{n}$,

and $C_n=\frac{1}{3}\sum_2^n\frac{1}{n^2}-\frac{1}{4}\sum_2^n\frac{1}{n^3}+\ \ldots\ +\frac{(-1)^k}{k+1}\sum_2^n\frac{1}{n^k}+\ \ldots\ ,$

a convergent series, since it is alternating and its kth term approaches zero as k increases indefinitely (Lemma C).

If now we substitute the value of u from (2) in (1) we obtain

$$n!=\frac{(n+1)^n}{2}\cdot e^{-n+\frac{1}{2}+\frac{1}{2}S_n-C_n} \tag{3}$$

and it remains only to examine the right member of this expression as n becomes very large. This involves some difficulties which are overcome by making the following transformations.

Formula (3) holds true for any integral value of n. Let us then replace n by $2n$. We then have

$$(2n)=!\frac{(2n+1)^{2n}}{2}\cdot e^{-2n+\frac{1}{2}+\frac{1}{2}S_{2n}-C_{2n}} \tag{4}$$

Now consider the identity

$$n!=\frac{2}{1}\cdot\frac{4}{3}\cdot\frac{6}{5}\cdot\ \ldots\ \cdot\frac{2n}{2n-1}\cdot\frac{(2n)!}{n!2^{2n}}. \tag{5}$$

On replacing in the expression on the right, $n!$ and $(2n)!$ by their values as given by (3) and (4) respectively, we obtain

$$n!=\left(\frac{2}{1}\cdot\frac{4}{3}\cdot\frac{6}{5}\cdot\ \ldots\ \cdot\frac{2n}{2n^{-1}}\right)\frac{[1+1/(2n)]^{2n}}{(1+1/n)^n}n^ne^{-n}e^{C_n-C_{2n}}e^{\frac{1}{2}(S_{2n}-S_n)}. \tag{6}$$

Now by (D) the first factor on the right approximates $\sqrt{\pi n}$ as n becomes very large, by (B) each of the expressions $(1+1/n)^n$ and $[1+1/(2n)]^{2n}$ approximates e, by (C) C_n-C_{2n} approximates zero, and finally

$$\begin{aligned}S_{2n}-S_n&=1+(\tfrac{1}{2})+(\tfrac{1}{3})+(\tfrac{1}{4})+(\tfrac{1}{5})+(\tfrac{1}{6})+\ \ldots\ +1/(2n-1)+1/(2n)\\&\quad -1\qquad\quad -\tfrac{1}{2}\qquad\quad -\tfrac{1}{3}\qquad\qquad -1/n\\&=1-(\tfrac{1}{2})+(\tfrac{1}{3})-(\tfrac{1}{4})+(\tfrac{1}{5})-(\tfrac{1}{6})+\ \ldots\ +1/(2n-1)-\tfrac{1}{2}n\end{aligned}$$

which approximates $\log 2$ as is seen by putting $x=1$ in (A). Hence it follows that when n is large we have approximately

$$n!=\sqrt{\pi n}\cdot n^n\cdot e^{-n}\cdot e^0\cdot e^{\frac{1}{2}\log 2}=e^{-n}n^n\sqrt{2\pi n}.$$

APPARATUS TO FACILITATE THE CALCULATION OF THE MOMENTS OF A DISTRIBUTION

By Thomas N. Jenkins, *Columbia University*

Devices to aid calculation of standard deviations have been described by Toops [1] and Croxton.[2] The device of Toops consisted of a series of circular brass checks strung on piano wires stretched between two parallel steel bars. Starting from the bottom, each stands for one step-interval. On each wire is hung 50 checks, each check standing for a frequency. The step-interval is written on one side of each check, and the F, FD, and FD^2 values are written on the other side. In actual use, "one merely pulls forward on the wire a brass check whenever one wishes to make a tally mark corresponding to a frequency of the distribution." This device saves the trouble of calculating the FD and FD^2 values for each step, and after the distribution is finished, the various values can be added, and the mean and standard deviation found by substitution in the usual formulae.

In one of the devices described by Croxton, the checks consist of rectangular slips arranged in piles and fastened to a common baseboard by "an inverted 'U'-shaped rod or staple passing through a hole in the left of each slip." This device possesses two advantages over that of Toops: First, it is easier to construct; second, when the operation is finished, the F, FD, and FD^2 values are arranged in vertical columns — a condition which facilitates the work of summation. In the other device designed by Croxton, the "piles of slips are replaced by a series of disks, the data being placed radially on them so that any frequency for each group may be brought into line by rotating the discs." This mechanism can be operated faster than the other but for large groups and large frequencies, its vertical dimensions make it an unwieldy tool.

The apparatus described in this paper was developed independently a few years ago to assist in the computation of the moments of grouped frequency distributions of reaction-times. The construction of this apparatus can be gathered from the accompanying photograph. D is a drawing board fastened to a table by means of a light wooden frame. The frame is adjustable so that the board can be tilted at different angles of inclination. The T-square shown in the picture

[1] Herbert A. Toops, "Two devices for Aiding Calculation," *Journal of Experimental Psychology*, Vol. 9, pp. 60–66.

[2] F. E. Croxton, "An Apparatus to Assist in the Calculation of the Standard Deviation of a Grouped Frequency Distribution," this Journal, Vol. 20, pp. 532–36.

at T has a spring clip and slider attached to the head so that it will remain in position on the board without being held by the hand. Along the length of the blade is a sliding pointer P. Fastened along the top edge of the blade is a strip of paper S on which the step-deviations are numbered from left to right. L is a spacing device made by the Lineatime Manufacturing Company of Rochester, New York. F is a frequency table of FD^2's. The frequencies are indicated in a vertical

TABLE I

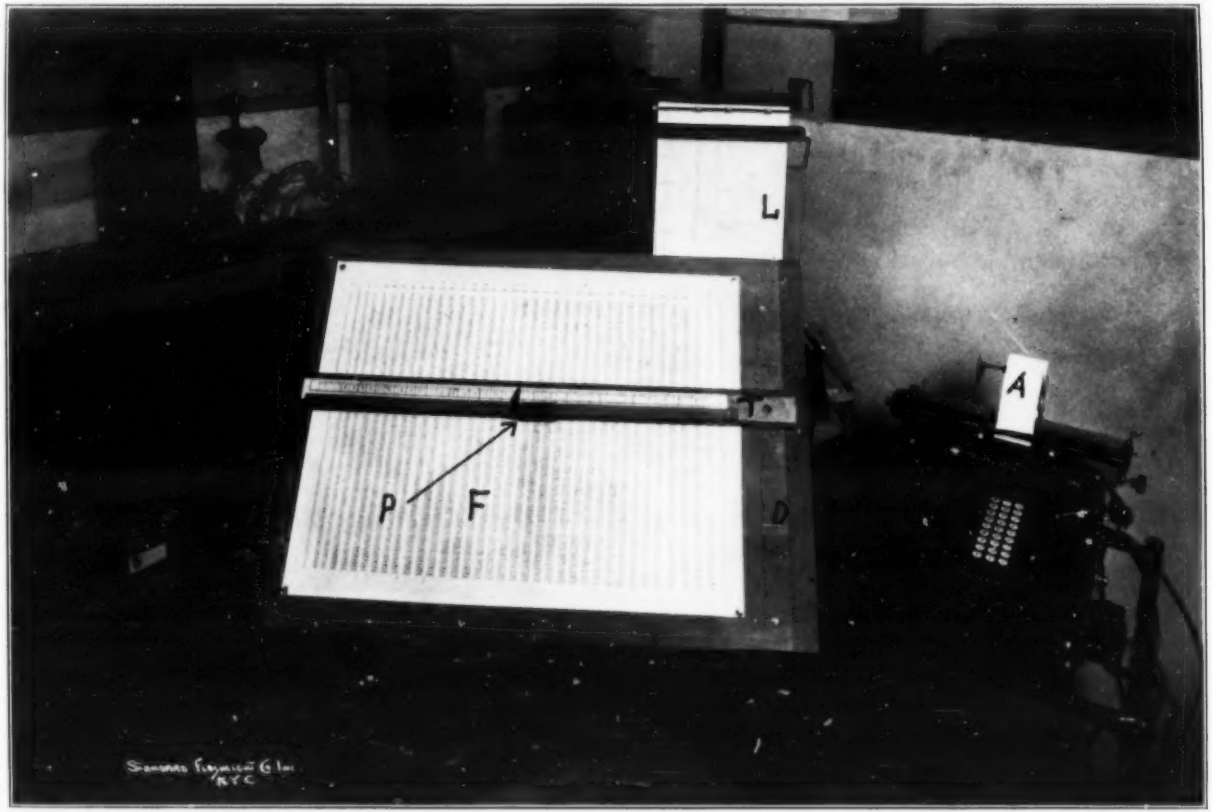
X	2	3	4	5	6	7	8	9	11	12
X^3	8	27	64	125	216	343	512	729	1331	1728
F										
2	16	54	128	250	432	686	1024	1458	2662	3456
3	24	81	192	375	648	1029	1536	2187	3993	5184
4	32	108	256	500	864	1372	2048	2916	5324	6912
5	40	135	320	625	1080	1715	2560	3645	6655	8640
6	48	162	384	750	1296	2058	3072	4374	7986	10368
7	56	189	448	875	1512	2401	3584	5103	9317	12096
8	64	216	512	1000	1728	2744	4096	5832	10648	13824
9	72	243	576	1125	1944	3087	4608	6561	11979	15552
10	80	270	640	1250	2160	3430	5120	6290	13310	17280
11	88	297	704	1375	2376	3773	5632	8019	14641	19008
12	96	324	768	1500	2592	4116	6144	8748	15972	20736
13	104	351	832	1625	2808	4459	6656	9477	17303	22464
14	112	378	896	1750	3024	4802	7168	10206	18634	24192
15	120	405	960	1875	3240	5145	7680	10935	19965	25920
16	128	432	1024	2000	3456	5488	8192	11664	21296	27648
17	136	459	1088	2125	3672	5831	8704	12393	22627	29376
18	144	486	1152	2250	3888	6174	9216	13122	23958	31104
19	152	513	1216	2375	4104	6517	9728	13851	25289	32832
20	160	540	1280	2500	4320	6860	10240	14580	26620	34560
21	168	567	1344	2625	4536	7203	10752	15309	27951	36288
22	176	594	1408	2750	4752	7546	11264	16038	29282	38016
23	184	621	1472	2875	4968	7889	11776	16767	30613	39744
24	192	648	1536	3000	5184	8232	12288	17496	31944	41472
25	200	675	1600	3125	5400	8575	12800	18225	33275	43200
26	208	702	1664	3250	5616	8918	13312	18954	34606	44928
27	216	729	1728	3375	5832	9261	13824	19683	35937	46656
28	224	756	1792	3500	6048	9604	14336	20412	37268	48384
29	232	783	1856	3625	6264	9947	14848	21141	38599	50112
30	240	810	1920	3750	6480	10290	15360	21870	39930	51840
31	248	837	1984	3875	6696	10633	15872	22599	41261	53568
32	256	864	2048	4000	6912	10976	16384	23328	42592	55296
33	264	891	2112	4125	7128	11319	16896	24057	43923	57024
34	272	918	2176	4250	7344	11662	17408	24786	45254	58752
35	280	945	2240	4375	7560	12005	17920	25515	46585	60480
36	288	972	2304	4500	7776	12348	18432	26244	47916	62208
37	296	999	2368	5625	7992	12691	18944	25973	49247	63936
38	304	1026	2432	4750	8208	13034	19456	27702	50578	65664
39	312	1053	2496	4875	8424	13377	19968	28431	51909	67392
40	320	1080	2560	5000	8640	13720	20480	29160	53240	69120
41	328	1107	2624	5125	8856	14063	20992	29889	54571	70848
42	336	1134	2688	5250	9072	14406	21504	30618	55902	72576
43	344	1161	2752	5375	9288	14749	22016	31347	57233	74304
44	352	1188	2816	5500	9504	15092	22528	32076	58564	76032
45	360	1215	2880	5625	9720	15435	23040	32805	59895	77760
46	368	1242	2944	5750	9936	15778	23552	33534	61226	79488
47	376	1269	3008	5875	10152	16121	24064	34263	62557	81216
48	384	1296	3072	6000	10368	16464	24576	34992	63888	82944
49	392	1323	3136	6125	10584	16807	25088	35721	65219	84672

column at the left edge of the table, and the step-deviations are printed at the head of each column of figures at the top of the table.

The operation of this set-up is as follows: The distributions to be handled are first clamped in the lineatime at L. Then a frequency table for the moment to be calculated is fastened to the drawing board with four thumb tacks in such a position that the horizontal rows of figures are parallel with the blade of the T-square when the head is pressing firmly against the edge of the board. Everything is now ready for actual computation. To illustrate, suppose we take an assumed mean at the midpoint of the first group of the distribution. The pointer P is moved over to the left under the step marked (2). The head of the T-square is moved up or down to the desired frequency, and the number opposite the pointer is recorded on the adding machine at A—the pointer P being moved one step to the right at each operation. The lever H is then depressed to bring the next frequency into view on the lineatime, and the process is repeated. The spacing device is not necessary, but it is very useful to eliminate the possibility of space errors in reading the frequencies and step-deviations of the distributions. The sums are recorded on each sheet before turning over (on lineatime) to another distribution. With a little practice, an operator can read off the values from the table as fast as an expert tabulator can record them on an adding machine. The writer with the help of an assistant found the averages and *SD*'s for 214 reaction-time distributions containing from 200 to 400 cases each in 16 hours. In this time each distribution was handled twice to check the results for errors.

Five frequency tables have been constructed for use in this machine. Three of these are for the second, third, and fourth moments around the arithmetic mean. A sample of the table for the calculation of the third moment about the mean is shown in Table I. The other two are for the first and second logarithmic moments around the logarithmic mean to be used in connection with a short method for handling logarithmic distributions.[3] With this machine, the moments around the logarithmic mean can be computed with a speed which is comparable to that for finding the moments around the arithmetic mean. The writer with the help of an assistant calculated the first and second unit moments of 54 distributions of maze data in three hours and twenty minutes.

[3] A description of this short method with accompanying tables will soon be ready for publication.


L
A
P
F
D

UNIV. OF MICH.

AVERAGE VARIATIONS IN YIELDS, ACREAGE AND PRODUCTION

By L. J. Norton

The relative importance of acreage and yield in determining variations in the total production of crops is of interest in connection with many agricultural problems. To what extent have farmers changed acreages in response to economic or weather conditions? What changes in acreage are equivalent to average change in yield per acre due to weather conditions? These questions are especially significant in considering proposed policies for stabilizing production and prices of farm products.

Average variation in acreage, yield per acre and total production from one year to the next for important crops are shown in the following table:

Crop	Period	Percentage of variation from previous year in		
		Acreage	Yield per acre	Total production
		Per cent	Per cent	Per cent
Corn	1900–26	2.6	13.0	14.5
Cotton	1900–26	8.0	11.2	16.9
Flaxseed	1900–26	20.3	20.6	29.0
Tame Hay	1900–26	2.5	8.2	10.0
Oats	1900–26	3.2	17.0	17.6
Potatoes	1900–26	4.9	17.1	19.7
Tobacco	1900–26	10.1	5.0	10.0
Wheat, winter				
Harvested	1910–26	14.9	9.5	18.1
Planted	1910–26	7.3	16.5	18.1
Wheat, spring	1910–26	9.4	32.6	30.7

The following may be noted:

1. Variations in the production of the three crops used chiefly for feed—corn, hay and oats—were chiefly due to variations in yield per acre, year to year fluctuation in acreage being negligible. Variations in acreage, in fact, largely reflected the trend or tendency for the acreage to increase over a period of years. Variations in yield per acre averaged in the case of tame hay 8 per cent, of corn 13 per cent, and of oats 17 per cent. The weather is a dominating influence in the production of these three crops.

2. Tobacco was at the other extreme, fluctuations in acreage being larger than in yield per acre. Considerable adjustment seems to be made from year to year in acreage and because of the quality of the land

on which it is grown, the care taken in its production, or a lack of sensitiveness to the variations in weather conditions in its normal area of production, the yield per acre varies but little.

3. Potatoes, another intensively cultivated crop grown on a small acreage like tobacco, resembled corn more than tobacco in respect to relative acreage and yield variations, the variations in yield averaging 17 per cent and the variations in acreage 5 per cent.

4. Cotton acreages varied on the average 8 per cent and yields 11 per cent, indicating comparatively large adjustments in acreage and relatively small fluctuations in yield. The boll weevil has, of course, increased variations in yield.

5. Flaxseed acreages and yields were both extremely variable, the fluctuations in both averaging about 20 per cent and in the total production about 30 per cent, indicating the flexibility of cropping systems and variability of the weather in the Dakotas and adjoining states that produce our flax.

6. Yields of wheat, both winter and spring, fluctuate more than acreages when acreages planted are considered. Fluctuations in total spring wheat production average as large as those in flaxseed, about 30 per cent, and are due to variations in yield rather than to acreage, the variations in acreage averaging only 9 per cent compared to 33 per cent in yields. The problem in winter wheat is complicated by the fact that a part of the acreage is not harvested and may be used for some other crop. On the basis of harvested acreage the fluctuation in acreage is larger than in yield, but this is due in a large degree to the abandoned acreage because on the basis of planted acreages the variations in yield are more than twice as large as those in acreage, 16 per cent and 7 per cent respectively.

It is mathematically possible to compute coefficients of determination measuring the relative influence of acreage and yield in causing variations in production. However, the actual relationships are more clearly shown by the percentage variations presented above than by this more refined measure. For example, variations in acreage of spring wheat were on the average 30 per cent as large as those in production, while variations in yield per acre equaled those in total production. It is obvious that during some years when acreage was increased the yield per acre was smaller or vice versa. The effect of these interrelationships is to reduce the net influence of acreage variations as shown by a coefficient of determination. However, if there had been no variations in yield, variations in production would have been 30 per cent as large as they actually turned out to be. That yield per acre happened

to be such as to offset these acreage fluctuations does not eliminate the actual influence of acreage variations on total production.

The real significance of changes in acreage is obscured in the figures in the foregoing table because, by the method of calculation, many acreage changes might influence the percentage for only a single year, although they would influence the production every year thereafter.

For example, suppose that acreages for a series of years were 100, 110, 110, 111, 112, the method used in preparing the data in the table would show an average variation in production of about 3 per cent while the actual acreage during the last four years averaged 11 per cent larger than during the first year. It is this cumulative acreage, not the year to year fluctuations, that must be considered in any study of the influence of price on production.

The changes in acreage necessary to equal the average variations in yield per acre shown in the former table using acreages from 1921–25 as a base for computing total production would be as follows:

Corn	13,300,000	Potatoes	600,000
Cotton	4,200,000	Tobacco	80,000
Flaxseed	400,000	Wheat, winter	7,200,000
Tame hay	4,900,000	Wheat, spring	6,400,000
Oats	7,300,000		

To give a complete picture this analysis would need to be extended to measure the relative importance of yield and acreage changes in different sections of this country and in foreign competing countries. The data here presented deal only with the aggregate changes for the United States as a whole and point out some considerations that need to be kept in mind in making an analysis of the relative influence of acreage on production.

OCCUPATIONS OF THE MEXICAN-BORN POPULATION OF TEXAS, NEW MEXICO AND ARIZONA, 1900–1920

By Roden Fuller

The number of Mexicans living in the United States in 1910 was more than twice as great as in 1900 and in 1920 more than twice as great as in 1910. At the last census there were nearly 500,000 Mexicans in this country and if the increase has been as rapid since that time as it was before, the Mexicans in the United States are now not far from 750,000. Quota restrictions do not apply to Mexican immigrants as they do to immigrants from European countries, and Mexican industrial conditions since 1920 have been depressed. For these reasons the rate of Mexican increase in recent years has probably risen.

Of the Mexicans in the United States nearly seven-eighths are in the border states of Texas, New Mexico, Arizona and California. The following study has been made in the effort to discover into what occupations the Mexicans are going. Unfortunately, the occupational statistics do not classify the foreign born or the foreign-born white by country of birth. For that reason, the statistics used are for the foreign-born white and have been confined to the states of Texas, New Mexico and Arizona, where, in 1920, between 69 and 79 per cent of the foreign-born whites were Mexican. The percentages of Mexican-born to foreign-born whites of recent years in the three states chosen were:

	1900	1910	1920
Texas	40.2	52.5	69.8
New Mexico	50.4	52.6	69.7
Arizona	63.3	64.1	78.8

For each state the ten occupations were considered in which the greatest numbers of foreign-born whites were found. When the lists were compared the following nine occupations were found to be common to all: Agricultural laborers; Farmers, planters, and overseers; Laborers (not specified); Merchants and dealers (excluding wholesale); Stock raisers, herders, and drovers; Steam railroad employees; Miners and quarrymen; Servants, waiters and waitresses; Laundresses.

Changes in occupational designations from census to census were traced and provision for such changes made except in the case of "Lab-

orers (not specified)" which was disregarded as too vague for our purposes.

1. *Agricultural Laborers.*—In 1900 agricultural laborers formed the second largest group of foreign-born white workers in Texas, while in 1910 and 1920 they held first place. Where in 1900 foreign-born whites represented 5.1 per cent of the total agricultural laborers in Texas, they represented 6.6 per cent in 1910 and 10.8 per cent in 1920. The increase of 4.2 per cent from 1910 to 1920 probably indicates the influence of the great influx of Mexican labor during that decade. In New Mexico and Arizona agriculture plays a lesser rôle. In New Mexico agricultural laborers in 1900 represented only the fourth largest group of foreign-born white workers; in 1910, the third; and in 1920, the second. Their percentage dropped slightly between 1900 and 1910 and then increased at the 1920 enumeration. The agricultural expansion of Arizona, particularly between 1910 and 1920, proceeded more rapidly than that of New Mexico. While in 1900 the agricultural laborers ranked sixth among the foreign-born white occupational groups, they ranked third in 1910 and second in 1920.

2. *Farmers, Planters and Overseers.*—In two states, Texas and New Mexico, the importance of the foreign-born white "Farmers, planters and overseers" decreased between 1900 and 1920. In Arizona, although the foreign-born whites increased their percentage, the group itself suffered a sharp decline in rank as an outlet for the activities of foreign-born whites. Further, the increase in percentage strength within the group between 1900 and 1910 was due, not to an advance in number of foreign-born whites, but to a great falling off in the numbers of other population elements.

The figures seem to indicate that little of the influx of Mexicans between 1910 and 1920 flowed into this particular occupation.

3. *Merchants and Dealers (excluding wholesale).*—Wherever large numbers of foreign born are settled those services required by the modern community tend to be administered by members of the same nationality. In the case of the Mexicans, who possess little knowledge of English, small stores are run for their fellow-countrymen by Mexicans.

The immigration of Mexicans between 1900 and 1910 was not large enough to stimulate a growth of foreign-born white merchants and dealers as rapid as that among the representatives of other population elements; so that in 1910 the percentage of foreign-born whites in this occupation decreased. But with the rush of Mexican immigrants, 1910–1920, came an increase among the foreign-born white dealers, greater proportionately than among the other population elements.

4. *Stock Raisers, Herders and Drovers.*—From 1900 to 1920 the im-

portance of the foreign-born whites as stock raisers, herders and drovers in the Southwest dwindled. The greatest number employed in this occupation was located in Texas, and in Texas the actual numbers decreased from 2,968 in 1900 to 2,300 in 1920, while the percentage representation decreased from 18.2 to 11.5, and the ranking of the occupation among leading foreign-born white activities from sixth to tenth. Although in Arizona there was a slight increase in actual numbers from 1900 to 1920, the activity fell in rank from fifth to tenth place, and the representation of the foreign-born group in it decreased from 13.3 to 9.7 per cent.

5. *Steam Railroad Employees.*—In no occupation did the foreign-born white group show such great increases as in that of steam railroad employees. Considering New Mexico, Arizona and Texas as a whole, there was a great displacement of other population elements by the foreign-born white group during the decade 1900–1920. The Negroes lost most ground. Mexican laborers, living on a scale below that of their black competitors and rendered amenable to discipline by a tradition of peonage, as well as producing less race friction, were increasingly used on the railroads of the Southwest and the section gang of Mexicans became a familiar sight in that part of the United States.

6. *Miners and Quarrymen.*—Except in the cases of Arizona and New Mexico in 1900, the foreign-born whites have been a majority among the miners and quarrymen of Texas, New Mexico and Arizona during the three past censuses; in Texas, as coal mine operatives; in New Mexico and Arizona, as copper, gold and silver mine operatives. Throughout this period, miners and quarrymen represented the greatest single group of foreign-born white laborers in New Mexico and Arizona. In Texas, where mining is of minor importance, they ranked near the bottom of the ten groups.

7. *Servants, Waiters and Waitresses.*—The foreign-born white group occupies a more important position in the domestic service class in New Mexico and Arizona than it does in Texas, the weakness in Texas probably being due to the greater number of Negroes available as competitors. In all three states the strength of the foreign-born whites in this occupational group increased from 1900 to 1920.

The trend of the evidence thus reviewed points towards a gradual displacement of Negro labor by Mexican labor in the main occupations open to both throughout the three states to which the study has been of necessity confined.

GENERAL TABLE

PERCENTAGES FORMED BY FOREIGN-BORN WHITES OF TOTAL POPULATION OVER TEN YEARS OF AGE ENGAGED IN CERTAIN SELECTED OCCUPATIONS, NEW MEXICO, ARIZONA AND TEXAS, 1900-20

Occupational designation	New Mexico			Arizona			Texas		
	1900	1910	1920	1900	1910	1920	1900	1910	1920
Agricultural laborers	7.1	6.4	9.9	16.0	23.0	18.9	5.1	6.6	10.8
Farmers, planters, and overseers	6.5	4.6	5.1	7.8	10.8	9.7	7.9	7.3	7.4
Laborers (not specified)	8.7	14.5	16.1	40.4	50.2	51.1	16.1	16.9	24.1
Merchants and dealers (excluding wholesale)	26.2	15.4	18.3	29.4	26.3	31.1	17.1	15.7	17.5
Stockraisers, herders and drovers	3.4	5.1	3.9	13.3	13.1	9.7	18.2	15.4	11.5
Steam railroad employees	15.1	37.8	32.6	37.5	78.3	81.0	14.7	25.8	48.0
Miners and quarrymen	39.8	58.2	56.2	41.5	62.9	62.7	62.8	61.0	63.8
Carpenters and joiners	15.4	10.2	13.5	....	....	....	14.0	9.7	10.8
Servants, waiters and waitresses	11.1	11.1	14.0	18.9	29.1	33.0	9.3	9.2	12.3
Laundresses	7.7	4.1	19.9	48.0	43.9	56.7	6.4	5.9	9.3
Copper workers	....	....	....	56.2	84.8	75.7	....	....	....

PRICE TRENDS AND PROSPERITY

A dinner meeting of the American Statistical Association was held on Thursday evening, October 27, 1927, at the Town Hall Club, 123 West 43 Street. About three hundred persons were present. Dr. Edmund E. Day, President of the American Statistical Association, acted as Chairman.

The first address of the evening was given by Mr. Carl Snyder of the Federal Reserve Bank of New York City. He pointed out that, at present, the movement of most of the index numbers of wholesale prices is upward, the long decline apparently having been stopped. The upward movement is primarily due to the rise in the prices of food and farm products. Mr. Snyder warned his hearers that it is entirely unsafe to assume, as is so often done, that the course of wholesale prices is indicative of the movement of prices in general. He showed, for example, that, since 1922, index numbers both of retail prices and of the broad list of commodities of all types compiled by the Federal Reserve Bank of New York, have moved in a direction practically horizontal, despite the very considerable fluctuations which have been taking place in the wholesale price index. He felt that this difference in trend was largely the result of the influence of wage payments, for the evidence at hand indicates that highly fabricated articles showed little or no tendency to decline at the time that wholesale prices were moving rather sharply downward.

Mr. Snyder emphasized the fact that it is possible, by utilizing available data on check payments, to check the accuracy of the index-number series representing the General Price-Level computed by the Federal Reserve Bank of New York. Ever since 1918, the Federal Reserve Board has received weekly reports covering from 80 to 90 per cent of all checks drawn in the United States. The theoretical price-level, computed on the basis of check payments, corresponds very closely to the composite index of the General Price-Level made up by combining retail prices, wages, wholesale prices, security prices, real estate prices, etc.

Another fact brought out by the speaker was that there is no evidence to show that fluctuations in commodity prices at wholesale precede general price movements. He felt that wholesale prices should always be regarded merely as one group of prices in general, important for certain purposes, but not for others. When the course of wholesale prices departs from that of prices in general, it may be expected to return eventually to the level of the more representative index.

The next speaker was Professor David Friday. He brought out the fact that, in the past, most students of business have believed that there is close, direct correlation between the movements of commodity prices at wholesale and industrial profits. The events of the last two or three years have, however, been such as to indicate that this rule is certainly far from universal for, although prices at wholesale moved downward rather steadily through 1925 and 1926, both years appear to have been extremely prosperous for most corporations engaged in transportation and manufacturing.

He said, "I think the answer to the puzzle is to be found in the productive output. Take the Harvard Index of Output for Mining and Manufacturing,

which, to my mind, is one of the most important and helpful contributions to economic statistics that anybody has made in a century: You will see that we are today producing in manufacturing about 38 per cent more than we were in 1919. The same is true in mining. And I have come to this one conclusion: That, as far as prosperity depends upon profits—and that is the fulcrum of this argument about the relation of prices and prosperity—the situation depends much more upon the volume of that productive output than upon the price-level; and that our serious problems for the future, growing out of the decline in prices—I doubt if it will continue next year—and if Professor Mitchell will assure us tonight that he sees no evidence of coming depression or of crises but that next year will see the largest production that we have ever had (my friend Julius Hirsch of Berlin will tell you a good deal about it)—then I will predict tonight that next year will be the largest year of industrial profits that we have had in this country up to date."

The third speaker was Dr. Wesley C. Mitchell, Director of the National Bureau of Economic Research. In reply to a query put by a previous speaker concerning the probable future trend of prices, he said:

"I have, however, a very comfortable reason for pleading marked uncertainty concerning this subject, a reason that pleases me, because I think my inability to forecast, even with reference to the next five or ten years, what the trend of prices is going to be, is a reason that arises from faith in the slowly increasing intelligence of mankind. That sounds cryptic and possibly I don't make it very clear, shan't make it very clear, even when I am done; but I shall try to.

"When we forecast the future, we do it by projecting into the years to come our interpretation of what has happened in the past. The reason why these projections of ours so often go astray, is that certain factors of great importance in the situation as a whole keep changing. We have a very beautiful illustration of this difficulty of forecasting provided by that curious experience which Dr. Friday analyzed. It's true of course that, had anyone been able to forecast in 1923 that the general trend of non-agricultural prices would be a declining one for the next four or five years, almost everyone would have taken it for granted that the period was going to be one of great difficulty for business enterprises, a period of meager profits. A person who was thinking about that problem and looking over historical parallels might have found a few instances in which declining price trends had been accompanied by good business. In our own history as a nation, we have had at least one rather striking parallel with this recent episode; the period of activity which preceded the great panic of 1873 was a period during which prices scarcely rose. They fluctuated about an almost horizontal level for several years before we had one of the most severe panics in our history. And, in the English record, you can find at least one other case that is described in the interesting statistical dissertation of Dr. Silverly. But cases of that sort are rare."

Dr. Mitchell pointed out further that the person who was guided in his expectations by a projection of past experience had ample ground for expecting inflationary development in the United States during the last few years because of the enormous gold stock in the vaults of the Federal Reserve Banks, yet nothing of this kind occurred.

No one can tell what will happen, but present indications are that the gold stocks of the world will not increase materially during the next few years. The chances are that Europe, in the near future, will demand more gold as, industrially, it is recovering from the difficulties of the post-war period. These forces point to a decline in wholesale prices.

The speaker stated, however, that he had high hopes that the great commercial nations of the world "will develop the intellectual initiative and courage which is presumably all that is required to introduce such changes in human practices as will prevent forces, which if left to themselves would probably bring about a price decline, from bringing upon us that event."

He continued, "In the past, I think a great many economists have felt that there was something fine about the gold standard because it left the course of prices to be determined by natural events. It could not be manipulated. I am inclined to think that, if there is anything fine about the gold standard for human purposes, it is because it is an instrument which can be manipulated by men, because it is something which can be controlled to work out the sort of results that we desire.

"How that can be brought about, I have no means of knowing. It's a subject concerning which many different opinions prevail. My one faith there, is faith in the continued exercise of human ingenuity, the continued willingness of men to experiment courageously, to keep a record of what they do, and try to learn as they go along from the results of whatever innovations they effect—provided that the world at large faces this issue in a scientific spirit, with a determination to control the monetary policy in the public interest to achieve certain definitely recognized objects."

Professor E. W. Kemmerer of Princeton University spoke next. He began his remarks by discussing the outlook for gold production during the years immediately ahead. He said it was only a short time ago that the prediction was freely made that the production of gold would not rise to the pre-war rate before 1930, at the earliest. As a matter of fact, however, this rate has already been exceeded in South Africa. Two years ago, the speaker found there conflicting ideas as to the prospects of future gold production. Mines were being forced deeper than people believed possible a decade since, and grades of ore are now being worked which were wholly worthless in the old days. Professor Kemmerer thought that it was quite unlikely that the increase in the gold production of the world was destined to stop in the near future.

He said, further, "Of course, in the use of gold, we have realized remarkable economies in recent years. The legal reserve requirements, for instance, of the banks in this country are only about one-fifth of what they were before the initiation of the Federal Reserve system; and, with the tendency in advanced countries to use less and less gold in circulation, and the development of our credit system and our banking machinery and the inter-relation of the central banks with the world, and the coöperation of these banks, there is likely to be very remarkable economy in the use of gold. There is the second factor looking toward rising prices.

"Then, we have in this country an enormous supply of gold that is practically

impounded. We can use it; we are not using it, and, sooner or later, as the world needs it, it will probably be fed out, and fed out without any particular sacrifice. That is another factor that will tend to push up prices."

Professor Kemmerer stated that, of course, he realized the growing rate of physical production in the United States and some other countries, would increase the demand for gold and tend to force the price-level downward. Another influence in the same direction would be the fact that some of the backward countries are going on a gold standard for the first time. He pointed out, however, that, from 1897 to 1913, the wholesale price-level increased at the rate of 2.9 per cent on an arithmetical basis, or 2.4 per cent on a geometric basis, and that, had the growth proceeded at this same gradient, it would have been today at a point not far from its present location.

Should prices go up very much, this movement would, of course, tend to check the output of gold. The speaker's general view was that, owing to the fact that we are controlling the price-level much more than ever before, we are quite likely to have a fairly stable price-level for the next few years.

The last speaker of the evening was Colonel Leonard P. Ayres of the Cleveland Trust Company. He cited examples showing that, during times of rising commodity prices, everyone holding commodities gains a profit merely by waiting. In such cases, the success of the business man, particularly of the industrialist, depends rather upon his skill as a prophet than upon his ability as a producer. During the past three years we have had a period of decline in wholesale prices of non-agricultural products. This has intensified competition and has favored the larger firms. Colonel Ayres felt, however, that it was unsafe to assume that firms are efficient because they are large. In his view, firms are large because they have been and are efficient.

This period of falling prices has been characterized by a regime of hand-to-mouth buying. To illustrate how far this has been carried, the speaker stated that, when lake navigation opened in the spring in Cleveland, it was found necessary to furnish the stevedores on the docks with heavy gloves and give them new handling tools because the automobile parts that they were moving were so hot, they could not load them in the ordinary way. Such hand-to-mouth buying makes for low inventories; high wages for labor, but not for strong competition for labor; speculation in real estate and security markets, but not speculation in debts.

Colonel Ayres stated that, "Success has depended and does depend now on efficient producing and effective selling, rather than on skill in forecasting the industry; and the industry itself is not prospering as it should." Furthermore, he was inclined to believe it rather probable that we are going to enter 1928 with business rather more depressed than it is now and that there will be some reaction in the spring, largely because of the coming back into production of the Ford Company. He anticipated for 1928 a disturbed summer with a turbulent Congress and some of that doubt which always gets injected into the business situation when a campaign is on. However, he looked for the latter part of 1928 to be a period of increasingly good business with large profits.

Should the upward trend in wholesale prices continue, business will boom, at

least for a time, and hand-to-mouth buying will go by the board. The speaker's observation was that "All business men love to speculate. They would rather speculate in the goods that they deal in, because they know those better than they do anything else. If they can't do that, as they have not been able to recently, they will speculate in real estate and the stock market. But if and when commodity prices turn up once more, I think that these prudent procedures of business that we have had recently will tend to disappear and that we will see a recurrence of speculation in the open commodity markets and in the inventories."

W. I. K.

NEW METHODS OF MEASURING PUBLIC HEALTH PROGRESS

A dinner meeting of the American Statistical Association was held on the evening of November 30, 1927, at the Town Hall Club, 123 West 43 Street, New York City. About forty-eight persons were present. The presiding officer was Dr. Robert E. Chaddock of Columbia University.

The program was opened by Dr. Edgar Sydenstricker, Chief Statistician of the United States Public Health Service. He spoke on "Methods of Measuring Results of Public Health Work with Especial Reference to the Prevention of Diphtheria." He felt that it would be extremely difficult to measure the results of public health effort in general, and conceived the purpose of the meeting to be "to find out whether or not, or how far, a specific social effort, which seeks to attain a specific objective by a specific method, actually accomplishes the result it is designed to accomplish." He warned against the danger of attempting to "put over" new public health measures before it is conclusively proved that they are beneficial or do not yield deleterious by-products worse than the evils which they are intended to remedy.

He said, "The recent development of public health as a science may be said to be along two important lines, at least: (1) In the determination of the efficacy of a method of prevention supplied by another science (*e. g.* bacteriology) in preventing a disease under actual conditions; (2) the discovery, by scientific procedure, of new methods of disease prevention, which implies the collection of data, the forming of hypotheses, the testing of these hypotheses by observations and even trials according to the principles of experimentation, etc. The measurement of results is thus an essential part of the science of public health."

He chose to confine his remarks, in the main, to a discussion of one particular phase of the modern health program, namely, the attempt to measure the effect of various devices used for combating diphtheria, which is just now one of the major immediate objectives of public health efforts. He stated that the present evidence is that, while diphtheria may be transmitted from a person suffering from the disease to another, most cases of diphtheria are probably transmitted by carriers, that is, by persons affected with the bacilli, but not suffering any ill effects therefrom. In cities, and in some rural districts, a large proportion of adults are immune to the disease because they have already been infected. Immunity to the disease varies with age and, apparently, to some degree, with race. The virulence of the diphtheria organism varies with climatological conditions and, at some times, it seems to be more intense than at others.

Dr. Sydenstricker continued: "The preventative efforts that are commonly practiced are: (1) Active immunization of susceptible persons, who are chiefly under 10 years of age, by the administration of toxin-antitoxin; (2) prevention of contact with sources of infection, especially cases from whom massive dosage may be received; (3) passive immunization by administration of antitoxin to cases."

"Previous to about 1900, isolation was practiced, more or less, as well as measures which are now regarded ineffective, such as fumigation. Since then, the use of antitoxin has been fairly general, although by no means as universally applied as it should have been. More recently, in the last ten years, active immunization has become more and more widely practiced and is being rigorously undertaken by public health agencies at the present time." Dr. Sydenstricker furthermore stated that this immunization has been aided greatly by the Schick test which shows whether or not individuals are immune. He also pointed out that the problem at hand was to measure the effect of new public health measures upon the rate of mortality from diphtheria. It has been asserted that diminution in the mortality rate during recent years may be ascribed to the use of the Schick test and the immunization of susceptible persons. Dr. Sydenstricker presented a chart recording the mortality from diphtheria and croup in New York City, between 1868 and 1926. It showed clearly that, between 1875 and 1890, there was a steady but slow downward trend in the mortality rate. With the adoption, in the early 90's, of the antitoxin treatment, the trend turned sharply downward and continued along an approximately straight line until about 1922, the death-rate having fallen during this period from about 70 to 16 per hundred thousand. Since 1922, the work of immunization by toxin-antitoxin has been going on and the trend has turned down more sharply than ever, the mortality rate now being reduced to only about 8 per hundred thousand.

It is not, however, legitimate to ascribe the entire decline from 16 to 8 to the toxin-antitoxin treatment for, if the preceding trend had been continued, the present rate would have been only about 12. Therefore, we may legitimately assume that the immunization policy has cut the rate from something like 12 to something like 8.

This illustration was presented as a sample of the care which must be taken in measuring the effects of public health policies.

The next speaker was Dr. Emma Winslow, Director of Research, Commonwealth Fund, Child Health Program. Dr. Winslow pointed out that the Commonwealth Fund has been interested in planning programs for child health improvement and in evaluating the results obtained by child health work now being done. To measure results, it is usually necessary to make use of records actually at hand. After a considerable amount of experimentation, it has been decided that the three best methods of measuring results are:

1. Per capita expenditures for the local support of health work.
2. The score card method of appraisal of public health work devised by the American Public Health Association.
3. The actual record showing the status of child health.

Manifestly, the first two methods show only the amount of interest in public health work in the community, and not the effects of such work upon child health. It is difficult to ascertain just how effective the work has been. The areas covered by intensive child health work are too limited to make the vital statistics obtainable therefrom, dependable. One test is to measure the heights and weights of the children of these areas and note the changes taking place after the installation of the child health programs. These records may be compared with those from similar localities in which no definite program has been adopted. Comparisons of physical defects may also be made in a similar manner. Records may also be obtained of the number of physical defects corrected, as shown by medical, nursing, and school records.

It appears possible to develop, from available records, measurements which may be standardized and which will actually show the progress due to the child health programs.

The third speaker of the evening was Dr. Raymond Franzen, Research Director of the School Health Study of the American Child Health Association. Dr. Franzen dealt with the "Biometric Approach to Child Health Problems." He emphasized the fact that one must be extremely careful, in dealing with available records, if he is to be sure just what the records mean, and to determine just what bearing they have upon the problem. Manifestly, the problem needs to be stated with great care. The data used should be objective and should eliminate, in so far as possible, questions of individual judgment. In most cases, it is necessary to deal with sample data, and, in such instances, it is, of course, imperative that the sample be representative. Great care must be taken to test the reasonableness of all assumptions made. The very best statistical methods are required in order to ascertain whether or not there is any real relationship between the supposed cause and the supposed effect.

The program was closed by Dr. W. F. Walker of the American Public Health Association, Field Director of the Committee on Administrative Practice. He discussed the "Appraisal Form as a Measure of Public Health Progress." He showed that, since the installation of public health work, there has been a demand for some rather simple and direct method of comparing different communities on the basis of their stage of advancement in public health work. For this purpose, a number of different criteria have been suggested. The most frequently used indices are death-rates and per capita expenditures. Mortality rates are not comparable unless they are adjusted for age, sex, and racial differences in the population. Furthermore, there is a considerable and variable lag between the initiation of proven public health procedures and their effect upon morbidity and mortality.

The difficulty of rating different communities on the basis of per capita expenditures for health purposes is that, as a rule, the only records available are those of official agencies, while the expenditures of voluntary health agencies may be very considerable. Furthermore, these records give no weight to the sufficiency of local medical practice.

Because of the difficulties just mentioned, it has been felt necessary to devise methods of rating which are more dependable than either mortality rates or per

capita expenditures. The approved rating method, at the present time, is that of making for each community a score card appraising its various health activities. As early as 1903, Woodward introduced a dairy score card covering the procedures and details necessary to produce safe and clean milk. In 1914, Dr. Chapin, in his work, made use of an appraisal form for studying state health activities. He assigned relative values to various activities, but was unable to find any standard by which to judge the adequacy of a given service.

"In 1917, Doctors Meader and Brooks applied to cities of New York State the first score form for purely city health activities, which consisted of 71 criteria, under 13 general headings, with 1,000 points allocated among these. Only 9 of these 71 criteria had specific directions for scaling down the points allowed."

In 1920, the American Public Health Association appointed a committee to prepare a schedule of appraisal for cities, and Dr. Chapin, in 1923, outlined the first draft of the appraisal form now in use. In 1924, the American Child Health Association also devised an appraisal form. Since that date, the two Associations have combined efforts to revise the appraisal form and have tested it in a large number of cities. This form, as it stands, consists of:

1. A list of activities common to the health work of most cities.
2. A schedule of relative values of these various activities arrived at by a group judgment.
3. A series of standards for judging the adequacy of the service rendered in these activities.

For 76 of the 131 criteria used, there are definite provisions for scaling down the relative values which have been assigned as the performance of the cities recedes from the standards established. In the other 55 criteria, no specific directions for scaling are given. There has been no attempt to cover any activity exhaustively, the object being merely to pick certain representative services as samples, knowing that these will reveal the character of work throughout the department.

In order to ascertain whether or not the appraisal form gives any indication of the effectiveness of public health work in the given cities, persons fully acquainted with the public health problems of these cities were asked to array them in the order of the excellence of that work. These rankings showed a relatively high degree of correlation with the results obtained by the appraisal form. There is also quite a definite tendency for the standards of excellence shown by the appraisal form, to vary directly with the per capita expenditures of the public health work of various cities. Dr. Clarence E. Ridley, of the School of Citizenship and Public Affairs of Syracuse University, has also submitted the appraisal form to critical analysis and, by use of certain mortality rates, judges that there is a definite correlation between the degree of public health activity and these mortality rates.

The most common use of the appraisal form is by local health officers to compare the status of their own accomplishments, with generally accepted standards, and to show the progress of local health work from year to year. The ultimate test of reliability, as a measure of actual health progress, will come when it is

shown that the scores, as given by the appraisal form, are closely correlated with morbidity and mortality rates.

The meeting adjourned.

W. I. K.

PROGRESS OF WORK IN THE CENSUS BUREAU

THE CENSUS OF MANUFACTURES

Practically all the 1925 statistics of manufactures have been published in the form of printed pamphlets, the contents of which will later be assembled into a single bound volume. The last of the pamphlet reports was issued on November 21, 1927. The volume is now in process of printing, and will, it is expected, be issued in the near future.

The blank schedules for the manufactures census of 1927 were mailed during the week beginning January 9 to a list of about 325,000 establishments. Experience shows, however, that probably not more than two-thirds or three-fourths of these "prospects" will qualify for the census. The number of "special schedules"—*i.e.*, schedules especially adapted for canvassing certain industries or small groups of closely related industries—to be used at this census is 153, as against 89 used at the census for 1925. These 153 schedules cover 226 of the 340 industries embraced in the census classification. The number covered by the special schedules in use at the 1925 census was 121. The industries for which no special schedules are provided are canvassed with the general schedule.

Each special schedule calls specifically for data as to the values, and in many cases the quantities also, of the several classes of products made in the industry. In addition, some of them call for data as to quantities of the principal materials consumed, and a few carry inquiries as to equipment in use.

At the coming census, for the first time in eight years, the Bureau will collect separate data in regard to the cost of fuel and purchased power (as a single item). At preceding biennial censuses this item of cost has been included, without segregation, in "Cost of materials."

The amounts of anthracite and bituminous coal consumed will also be reported. This was done at the quinquennial censuses and at the biennial census for 1923 but not at the biennial censuses for 1921 and 1925.

COMPLETING THE 1925 CENSUS OF AGRICULTURE

The three main volumes of the census of agriculture, giving statistics by counties for the Northern, the Southern, and the Western states, respectively, each with a summary for the United States, have been published.

The first deficiency act, approved December 22, 1927, carried an appropriation of $55,000 for additional work on the census of agriculture. This was secured to enable the Bureau to make further compilations for which the former appropriations did not suffice. It provides for a number of additional runs of the punched cards on the tabulating machines. When the bulletins and volumes already issued were prepared it was possible to classify only a few items by tenure. The additional items now to be classified by tenure include pasture land, value of

buildings, number of farms by size, expenditures for feed, fertilizer, labor, and lumber, kind of road, distance to market, tractors, radios, acreage and production of specified crops, and number of specified classes of livestock. In addition the data on mortgage debt are being classified by size and value of the farm, and statistics are being compiled for a number of items which under former appropriations could be carried only as far as the completion of the machine tabulation. The Bureau hopes to publish a bulletin giving a summary of these statistics for the United States, by states. This will complete the 1925 census of agriculture.

FINANCIAL STATISTICS OF STATES AND CITIES

The summary of the 1926 report on financial statistics of cities having a population of over 30,000 was completed in December.

In January special press summaries were issued on expenditures for police departments, criminal courts, and correctional institutions, and on public school expenditures and indebtedness, for all of the 250 cities having over 30,000 population, 1926. Another special summary, showing the expenditures of the cities for conservation of health, sanitation, or promotion of cleanliness, charities, and hospitals, 1926, was released in February.

The inheritance tax and income tax portions of the *Digest of State Laws Relating to Taxation and Revenue, 1922*, have recently been brought up to date in the Bureau, and published in mimeographed form, by states, thus giving the recent state legislation on these subjects.

THE CENSUS OF DISTRIBUTION

The Bureau of the Census has now issued preliminary figures for all the cities covered by the recent census of distribution including Atlanta; Baltimore; Chicago; Denver; Fargo; Kansas City, Missouri; Providence, including Pawtucket and Cranston; San Francisco, inluding Alameda, Berkeley, and Oakland; Seattle; Springfield, Illinois; and Syracuse.

J. A. H.

SOME SUGGESTIONS FOR SECURING STANDARDIZED CRIMINAL STATISTICS

In its recently issued manual for compiling criminal statistics the United States Bureau of the Census rightly points out that "the lack of adequate criminal statistics has long been a weak point in the administration of criminal justice in the United States. In many cities and states police departments, courts, and penal institutions neither compile nor publish any criminal statistics whatsoever. Furthermore, where such statistics are published, they are not compiled in accordance with any standard plan."

If these conditions are to be overcome it is believed that state action will have to be secured. It is apparent, however, that if the states are to be brought in line and the inertia toward adequate statistical reporting by the various institutions, courts, etc., be overcome, considerable educational work will have to be done.

A strong impetus toward effective and unified state action might come from a National Conference of representatives of State Departments supervising penal institutions and agencies, called solely for the purpose of discussing the furtherance of standardized criminal statistics on a state-wide basis. To this conference should also be invited representatives of the various national organizations—public and private—who have a direct interest in criminal statistics and whose support would be needed in securing the desired action.

Preliminary to such a national conference on standardized statistics it would seem desirable to have the National Crime Commission, in conjunction with the United States Bureau of the Census and other national organizations interested in crime statistics, undertake a survey which should elicit the following information:

1. What penal institutions, courts, etc., are now required to make reports of a statistical nature to a central state agency.
2. What penal institutions, courts, etc., are not now required by law to make reports of a statistical nature to a central state agency.
3. What does the law specifically require penal institutions, courts, probation and parole agencies, etc., to report to the central state agency.

The agenda for this conference might be furnished by the results of this preliminary survey of the status of criminal statistics of the United States. The principal aim of this conference would be to outline a program for creating the necessary state machinery through which would be obtained standardized criminal statistics which would (a) render a comparable picture of crime conditions from state to state or community to community; and (b) fit into a national scheme of criminal statistics.

The program drawn up by the conference should act as a guide for the self-examination of the status of each state in regard to its criminal statistics and lead to a decision of what specific action seems to be required in the light of this program. Each state should be able to decide:

1. What changes or improvements are needed in its reporting system as at present organized, covering the reports of the penal institutions, courts, etc., made to the central agency of the state and the compilation of these reports by the central agency.
2. What other institutions besides those now reporting must be included in a complete crime reporting system and what seems the most desirable action to bring them into line.
3. If the present state laws governing the reporting of crime are either inadequate or antiquated what legislative action seems desirable.
4. How the reporting of the individual institutions, courts, etc., to the central state agency can be so organized as to permit the state to compile the data to meet its peculiar state requirements and at the same time have the data fit into the system now established by the United States Bureau of the Census or to be established by it in the future.

EMIL FRANKEL

New Jersey State Department of Institutions and Agencies

A STANDARD REPORTING SYSTEM FOR PUBLIC ACCIDENTS

Although industrial accident statistics have been collected for twenty years a recent critic [1] concludes that prevention efforts are still handicapped because "there is no uniformity in the collection and tabulation of accident data" and "as a consequence the data for particular industries in different states cannot be combined with significant results."

The seriousness of public accidents became apparent ten years later than that of industrial accidents, but steps have already been taken to avoid the present defects of industrial accident statistics by inducing states and communities to adopt a uniform system of public accident reporting.

The plan, known as the Standard Accident Reporting System,[2] was devised shortly after the first National Conference on Street and Highway Safety in 1924 by a group of statisticians, public safety officials and engineers acting as the Statistics Committee of the National Safety Council. Dr. Louis I. Dublin of the Metropolitan Life Insurance Company has been chairman of this committee.

The Standard System after only two years of promotional work is now used by the states of New York and North Carolina and by some sixty communities including Chicago, Detroit, and other important places, and covering in all more than twenty-five per cent of our national population. If present plans develop, the states of Pennsylvania, California and Oregon will also adopt this plan beginning January 1, 1928.

The adoption of the system, which follows closely the recommendations of the National Conference on Street and Highway Safety, has been hastened by the official endorsement of the International Association of Chiefs of Police, the Accident Committee of the Highway Research Board, and the American Association of State Highway Officials.

The system meets the needs of city departments and state motor vehicle bureaus, both where reports of motor vehicle accidents are compulsory and are made by the individuals involved in the accident and where only such accidents are reported as come to the attention of police officers. The percentage of all accidents which a reporting agency will receive evidently depends upon which of these situations exist and the trend seems to be definitely toward compulsory reporting.

The system involves the use of three forms—a primary report card, a monthly tally sheet, and a monthly summary sheet. The reporting card is so designed that much of the required information is given by check marks. The monthly summary sheet, compiled either by the use of the tally sheet or by mechanical tabulation presents three general kinds of information: (1) A general summary by type and age, showing fatal and non-fatal injuries and property damage accidents, by age groups and by types and sub-types of accidents; (2) a classification of types of motor vehicles involved; and (3) the circumstances of motor vehicle accidents.

[1] Robert Morse Woodbury in *Workers' Health and Safety, A Statistical Program.* The Macmillan Company, New York, 1927.

[2] *Public Accident Reporting*, obtainable from the National Safety Council, 108 East Ohio Street, Chicago, contains a complete explanation of the Standard Accident Reporting System.

These circumstances give information on the following points: (1) Actions of the driver at the time of the accident, including any violation of traffic rules; (2) the condition of the driver; (3) the sex of the driver; (4) the condition of the vehicle (information being given about defective brakes, etc.); (5) the pedestrian (where he was and what he was doing); (6) condition of the pedestrian (confused by traffic, etc.); (7) sex of pedestrian; (8) condition of the road; (9) weather conditions; (10) light conditions; (11) railroad crossing (whether guarded or not); (12) location of the accident (at intersection, bridge, etc.); (13) residence of the motorist (out of city, out of state, etc.).

An essential feature of the system is that all of the information asked for is objective in character and requires a minimum of judgment on the part of the reporting person and the statistical clerk. Questions are not asked as to which driver was to blame, whether the pedestrian was to blame, etc. The purpose of the plan is not to establish responsibility for the occurrence of accidents but to provide for the city or state department a record of all the important facts about accidents, enabling them to study prevention methods more intelligently.

It is true that this statistical plan will not solve the accident problem. No such plan will. The belief of those who inaugurated the system, however, was that the facts obtained would serve as a foundation for special studies, both statistical and engineering, which would develop practical remedies for the public accident situation. The plan is being recommended by the National Safety Council and by the states and cities that are using it, in the hope that public safety officials will not continue to be under the handicap that industrial safety men have had to face because of the lack of uniform comprehensive statistics.

R. L. Forney

National Safety Council, Chicago

MISCELLANEOUS NOTES

Meeting of the San Francisco Chapter.—The fifth dinner meeting of the San Francisco Chapter was held on Wednesday, December 14, 1927, at the Bellevue Hotel.

The topic for consideration was "The Problem of Collecting Statistical Data and the Means Adopted for It."

The discussion was opened by Mr. Allan Sproule of the Federal Reserve Bank of San Francisco, who spoke from experience of the Bank in securing information with respect to trade reporting.

Mr. Sproule emphasized that data must be sought from all available sources, and that data which are needed must be balanced against data which can be secured. There must be reciprocity in the exchange of information or at least in advising the coöperating parties of the results of the investigation. As a rule studies must be made of general and not special trade movements. Where the geographical factor is present weights must be assigned to areas to give consideration to data in accordance with their importance in the trade of the district. It is often possible to preserve the homogeneity of a series through samples of data. The best way to obtain coöperation is to secure a few leaders in the particular industry, after which the others will generally come into line.

A questionnaire must be simple and easily understood. It should ask for only the data actually needed, in the manner in which they are usually compiled, and after conferences with those engaged in the industry. Where delays in filling out the forms are frequent the coöperators must be followed up by telegram and letter. After the data are in the office they can be summarized by mechanical systems which save a great deal of labor and many mistakes. Data furnished must of course be kept strictly confidential.

Dr. H. E. Erdman of the University of California was the next speaker. He spoke from the viewpoint of one gathering agricultural data.

He stated that with agricultural data there was usually no problem in keeping the data confidential, although it was often advisable to do so. In connection with farm management the question of cost of production became important during price fixing operations at the time of the War. Surveys as to the cost of production or the farm labor income must be conducted by personal interviews, and usually the farmers are suspicious of such questions. Cost of production figures are particularly hard to secure for particular crops. At the present time the stress is being placed upon the income of the farm as a unit under the typical organization, including dairying, fruit raising, and other branches. Attempts are made to get production in terms of physical units as well as in money value and then to set up the ideal production schedule for the farmer. Cost figures are of little significance in individual farmer's operations.

The criticism of the agricultural colleges has been that they teach the farmers how to produce more but not how to market the extra production. Recently studies have been along the line of market analysis. They have fallen into several classifications such as:

1. Studying the steps in the process.
2. The split-up of the consumer's dollar.
3. Which is the least costly marketing route?

Jobbers will usually give no information whatever, even if they have it, which is often doubtful. The problem of joint costs in marketing also enters in. Most investigations have to study costs in connection with a particular organization going directly to the books with the aid of accountants. Recent studies have sought to

determine the margins between the steps in marketing, such as the margin between the wholesaler and the consumer. Another method is to study margins in the same organization over a period of years.

The discussion was closed by Professor A. H. Mowbray of the University of California, who spoke of securing data in connection with insurance problems. He first discussed the workmen's compensation data for use in making rates. The reports of accidents by states, wages, and causes are full of errors especially as to causes of the accidents. This is brought out in studies of accidents by industry or in relation to specific hazards.

In automobile liability rate making there are many modifying factors affecting accidents. However, by selecting four general types of cars, regional experience may be compared with national experience. In working up data for California an estimate of the number of insured cars was made from data furnished by the Underwriters Association, which included only part of the companies. The number of owners financially responsible for damages was estimated based upon the number of machines purchased on contract, with adjustment for those who preferred to purchase in this manner but who were able to pay cash. This estimate was checked by income tax returns showing the number of incomes of $2,000 or more in the state. In many cases estimates have to be made along the same lines where there is very little direct data available. One such study was relative to old age pensions in California.

The Los Angeles Chapter.—So far there have been three meetings held this winter. On September 28 the subject discussed was "The Fitting of Long Time Trends to Time Series." The paper presented by Mr. W. F. French emphasized the importance of choosing a proper time interval for the fitting of the long time trend. While it is desirable, whenever possible, to use a trend fit by mathematical means, this careful work can be vitiated by a careless choice of time interval. It was also brought out that in studying and discussing the economic data of a relatively small district, the use of a trend covering a relatively long period was probably not as desirable as the use of a shorter period providing, of course, that the fundamental requirements of a true trend are fulfilled. While the trend over a long period is invaluable as a guide to the statistician, the business man prefers to base his comparisons on a period within his memory.

At the November 8 meeting a paper was presented by Mr. J. C. Clendenin on "Correlation and Regression." Correlation and regression methods are mathematical measures of the relationship between different series of data. Regression relates to the nature of the relationship, while correlation measures the closeness of the relationship. These devices are mathematical, and very easily misused, so they must be used with care. For any given value of one variable, the regression line indicates the most likely value of the other. Regression may be either linear or curvilinear. Multiple correlation methods are means whereby the combined result of the movements of a number of related variables upon one other may be measured. The point was brought out that two lines of regression were possible when correlation was not perfect.

At the annual meeting on December 13 the following officers of the Los Angeles Chapter were elected for the year 1928: President, George G. Ellis; Vice President, A. T. Anderson; Secretary, Ira N. Frisbee; Treasurer, W. F. French; Editor, J. R. Douglas; Counsellors, H. E. Craig, J. L. Niceley.

The Columbus Chapter.—The Columbus Chapter of the Association has held three interesting meetings this season. The October meeting was held jointly with the

evening session of the Ohio Conference of Statisticians. Miss Edith M. Miller of the National Bank of Commerce, New York, spoke on the subject of "Prices, Credit, and Prosperity." Mr. Karl G. Karsten of the Karsten Statistical Laboratory, New Haven, spoke on the subject of "Does Business Forecast the Stock Market, or Does the Stock Market Forecast Business?"

In November, Dr. E. R. Hayhurst, Professor of Hygiene, the College of Medicine, Ohio State University, addressed the chapter on the subject "A Study of the Pulmonary Condition of Nine Hundred Sandstone Quarry Workers." In December, Dr. A. B. Wolfe, Professor of Economics, Ohio State University, spoke on the subject "Post-war Population Literature."

Meeting of the Chicago Chapter.—Forty-three members attended the first meeting of the year at the High-Noon Club, Tuesday evening, October 25. Professor Horace Secrist gave us a very interesting description of his investigations in the field of banking during the past two years.

In the social half-hour before dinner, we learned enough about what the others are doing to make all anxious to learn more. Two ways are open to gratify this desire.

The first way was suggested in the general discussion following Professor Secrist's talk—that is, to issue a directory of members of the Chicago Chapter. This the secretary, Mr. S. W. Russell, will be glad to do.

The second way is to have members describe some phase of their work at the meetings this winter. The information asked for under "Principal Use of Statistics" will help the Program Committee to arrange for meetings of this kind.

United States Bureau of Labor Statistics.—Since the publication of the last quarterly statement of the work of the Bureau of Labor Statistics, prepared for this JOURNAL and published in the December, 1927, issue, the Bureau has begun a study of the various fields of activity which have been entered by the trade-unions of the United States. The first article prepared from this investigation appeared in the *Monthly Labor Review* for January, 1928, and covered the provisions made by the unions for the care of sick, aged, and disabled members, and for dependents.

In connection with the Bureau's survey of retirement systems for public service employees in the United States, a study is also being made of systems in effect in certain foreign countries. The first article of the foreign series was published in the *Labor Review* for January and related to Great Britain and France.

Data concerning wages, hours, and working conditions in cotton gins, cotton compresses, and cotton-seed-oil mills are now being collected by field agents of the Bureau. The field work on the following investigations has been completed and the reports are in course of preparation: Apprenticeship in the building trades; wages and hours of labor in the slaughtering and meat-packing industry; wages and hours of labor in foundries and machine shops; productivity in the iron and steel industry; and settlements made for accidents to officers and crews on vessels of American registry. The report of the last-named study will show the methods by which the settlements were made, the amounts, and how the settlements actually made would compare with the probable settlements had seamen been included in the Federal longshoremen's and harbor workers' compensation act of March 4, 1927.

Statistical Developments in Canada.—Statisticians and economists have long been convinced that that part of a nation's economic life about which least is known is wholesale and retail trade. Perhaps the first attempt to secure a survey of what is involved in the trading activities of a modern nation was undertaken by the Dominion

Bureau of Statistics about 1924 in the census of trading establishments and their activities covering the year 1923. While the statistics collected were incomplete, owing to the failure of many small establishments to furnish the data required by the Bureau, it is, nevertheless, considered that the results are of considerable value, and justify the effort which has been made in their collection and compilation. According to a bulletin recently issued, the total capital investment in 66,814 establishments carrying on wholesale and retail trade in Canada was shown to be $1,580,123,723. The total purchases of these establishments were $2,321,078,297 and the total sales $3,030,663,185.

In recent years selected members of the staffs in economics of various Canadian universities and colleges have spent their summer vacations in research work at the Dominion Bureau of Statistics. Messrs. Kenneth W. Taylor of McMaster University, N. A. Robertson of the University of British Columbia, and W. Burton Hurd of Brandon College, carried on research work there during the past summer.

Activities of the New England Council.—The activities of the New England Council are proving to be a stimulating force in the development of statistical data about New England.

Probably the largest undertaking in this field, growing out of the Council's activities, is the marketing and industrial survey of New England, undertaken by the United States Department of Commerce at the instance of the Council. In its study of New England as a marketing area, the Department has made an intensive analysis of the buying power of the New England population, and has endeavored to break down the New England area into major and subsidiary trading areas. For this work the statistics commonly used in these studies have been collected, and in addition certain other indices which, it is said, have not hitherto been included in such investigations. The results of the 1925 Census of Manufactures are being incorporated into the Department's reports on its industrial survey, together with data obtained from inquiries addressed to some 13,000 manufacturers concerning their sales and production trends, markets, chief sources of material, etc. The Council is now releasing, for publication in New England daily newspapers, short abstracts of reports on more than 100 classes of New England industry, based on the Department of Commerce studies.

The Policyholders' Service Bureau of the Metropolitan Life Insurance Company is another instrumentality that is doing work in the New England field at the instance of the Council. The Bureau is just completing a study of the use of research methods in New England business and the results, the purpose being to bring out clearly the relation between the use of modern research and the profitable operation of a business enterprise. The study has consisted of locating specific examples of the successful application of research methods to business problems by New England firms in the fields of employment stabilization, production and marketing. One phase of this study was a search for firms and commercial organizations developing and using statistics of economics. The investigation will report on approximately twenty cases under this category.

Leaders in the wooden box industry of New England have been holding a series of conferences this year for consideration of data concerning the state of their industry secured for them in a survey conducted by the Harvard Forest under the sponsorship of the Council. Their data include figures from every New England manufacturer in the wooden box industry, concerning such matters as annual lumber consumption, annual sales, freight shipments, prices paid for lumber, capital investment, etc., etc. The study revealed what was expected, namely, that the industry is sick, and not

operating on a reasonable profit basis, and the manufacturers who have been sitting in conference on the facts as presented have been trying to formulate a constructive program for relief of the situation.

The future programs of the New England Council, it is understood, call for considerable attention to the further collection and dissemination of statistical data concerning the development, resources, and economic activity of the New England area.

Answer to Professor King

To the Editor of the Journal of the American Statistical Association:
Sir:

Professor King is doubtless right in his main statistical findings, as presented before the Association in Washington, but wrong in interpreting them and in his criticisms of my own work in this field (with which he shows himself to be unfamiliar).

His sampling is, I think, equivalent to weighting by base weights, thus creating a "weight bias" to offset the "type bias" of the arithmetic formula. I have myself called attention to this very sort of offsetting in *The Making of Index Numbers* (third Edition, page 527).

The reason for preferring the geometric to the arithmetic or harmonic type is not that it is intermediate between them but that it is self consistent while each of the others contradicts itself, when used backward and forward.

I hope to elaborate on this reply at an early time.

IRVING FISHER

A Correction

To the Editor of the Journal of the American Statistical Association:
Sir:

I am very grateful to you for calling my attention to an error which I made by attributing the statement cited on page 411, September, 1927, of this JOURNAL to Mr. John Hilton instead of to Mr. Norman Crump of London, England, who made this comment in discussing Mr. Wood's paper: "An Examination of Some Statistics Relating to the Wool Textile Industry" read before the Royal Statistical Society, London, on February 15, 1927.

I regret exceedingly that this error occurred.

R. VON HUHN

James A. Field

James A. Field, Professor of Economics at the University of Chicago, died in Boston on July 15, 1927. He was born in Milton, Massachusetts, May 26, 1880, the son of James A. and Caroline (Whitney) Field. After graduating from Harvard in 1903 he spent the next two years in graduate work at that university and the following year in study abroad. In 1906 he returned to Harvard as Instructor in Economics and remained there for two years. In 1908 he went to the University of Chicago as an Instructor and there was advanced through the successive academic ranks until he was made a Professor in 1918. During the years 1909–10 and 1911–17 he was Managing Editor of *The Journal of Political Economy*. In the World War he served as a special investigator for the Division of Statistics of the Council of National Defense in Washington and then as chief statistician of the American Shipping Mission to the Allied Maritime Transport Council in London. In 1914 he married Amy M. Walker of Chicago. His widow and two sons survive him.

Professor Field's courses at the University of Chicago were mainly devoted to the subjects of Statistics, Population Theory, and The Standard of Living. In the field of

Statistics he was especially interested in graphic methods of presentation and his article on "Some Advantages of the Logarithmic Scale in Statistical Diagrams," published in *The Journal of Political Economy* for October, 1917, was one of the first to call attention to the possibilities in the use of logarithmic paper and still provides what is probably the most thorough discussion of the topic. However, Professor Field's main interest lay in the subject of Population. He was actively concerned in the eugenics movement and his research was primarily directed towards producing a history of English population theory since Malthus. He had written two articles on this subject, but the most thorough deliberation and the highest ideals of scholarship and workmanship tended to limit his publications. Of books he was joint author of *Outlines of Economics*, 1910, and joint editor of *Materials for the Study of Elementary Economics*, 1913. Upon his students as well as upon his friends his keen and subtle mind, his meticulous attention to detail and finish, the scrupulous fairness and deliberation of his judgments, his high ideals and character, combined with a physique and personality of distinction left an impression which will long remain.

Personals

Mr. R. H. Coats, Dominion Statistician, is attending the Seventeenth Congress of the International Statistical Institute, of which he is a member, as representative of the Government of Canada in response to an official invitation from the Egyptian Government requesting, on behalf of the Director of the International Institute of Statistics, that the Government of Canada be represented at this Congress.

There have recently been some changes in the organization of the Division of Statistical Research of the Bureau of Foreign and Domestic Commerce, Department of Commerce, Washington, D. C. Dr. E. Dana Durand, the chief of the division, is in Europe on an extended leave of absence. Mr. Lawrence B. Mann is acting Chief in his absence. Mr. R. von Huhn has been placed in charge of the Foreign Statistics Section of the division and Mr. E. H. Omohundro of the United States Section. Dr. Helen M. Strong and Dr. Theodore Henckels continue in charge of the Geographic and Translation sections, respectively.

During the past summer, Professor J. T. Sellin and Mr. W. W. Weaver of the University of Pennsylvania completed a bibliographical survey of all historical, statistical, and other research material available in Pennsylvania for studies in criminology. This survey was a part of a compilation made by the national Social Science Research Council, and is to be published by them in the near future.

MEMBERS ADDED SINCE DECEMBER, 1927

Arnold, John K., Federal Trade Commission, Washington, D. C.
Breithut, Frederick E., City College of New York, New York, N. Y.
Boyle, John W., Investment Securities, 36 Wall Street, New York, N. Y.
Caplan, Irving, National Industrial Conference Board, 247 Park Avenue, New York, N. Y.
Camp, Dr. Chester C., Associate Professor of Mathematics, University of Nebraska, Lincoln, Neb.
Coggins, Paul P., Statistician, New Jersey Bell Telephone Company, Newark, N. J.
Cogswell, Edmund S., Consulting Actuary, 161 Devonshire Street, Boston, Mass.
Cornwell, William C., Analyst, J. S. Bache and Company, New York, N. Y.

Daugherty, Dr. Carroll R., Instructor of Geography, Wharton School, University of Pennsylvania, Philadelphia, Pa.

Davis, Dr. W. Lloyd, Professor of Economics, Akron University, Akron, Ohio.

Dodson, Loren R., Statistician, Conway Company, 100 East 42 Street, New York, N. Y.

Durst, Walter N., Investment Analyst, Alexander Hamilton Institute, New York, N. Y.

Eastman, Roger M., Statistician, American Telephone and Telegraph Company, New York, N. Y.

Edgerton, Harold A., Department of Psychology, Ohio State University, Columbus, Ohio.

Elliott, George A., Fellow, University of Chicago, Chicago, Ill.

Ely, Owen, Statistical Department, Dominick and Dominick, New York, N. Y.

Faymonville, Philip R., Collector of Manufacturing Statistics, United States Army, 39 Whitehall Street, New York, N. Y.

Foster, Thorne, 721 East 21 Street, Brooklyn, N. Y.

Gaudreau, Armand T., The Barrett Company, 40 Rector Street, New York, N. Y.

Gill, Corrington C., Economic Research, 2630 Adams Mill Road, Washington, D. C.

Girard, Richard A., Alexander Hamilton Institute, New York, N. Y.

Goldsmith, Samuel A., Director, Bureau of Jewish Social Research, 114 Fifth Avenue, New York, N. Y.

Graves, L. M., The Brookmire Economic Service, Inc., 570 Seventh Avenue, New York, N. Y.

Greenleaf, G. Elton, Assistant Statistician, Kelley Drayton and Converse, 40 Exchange Place, New York, N. Y.

Hodgkinson, William, Jr., Statistician, American Telephone and Telegraph Company, New York, N. Y.

Horwitz, James W., Assistant Professor of Statistics, Harvard University, Cambridge, Mass.

Jackson, Theodore A., Instructor in Economics and Psychology, Warrenton, Mo.

Jenss, Rachel M., Vital Statistician, Medical Division, Indian Bureau, Department of the Interior, Washington, D. C.

Jones, Alan M., Statistician, Barrett Company, General Manufacturing Department, 40 Rector Street, New York, N. Y.

Kiessling, Dr. Oscar E., Government Research, Coal and Coke Section, United States Bureau of Mines, Washington, D. C.

Kirk, Lester K., Bureau of Business Research, University of Detroit, Detroit, Mich.

Larrabee, William M., Statistician, American Telephone and Telegraph Company, 195 Broadway, New York, N. Y.

Light, Dr. G. H., University of Colorado, Boulder, Colo.

Lundberg, Dr. George A., Department of Sociology, University of Pittsburgh, Pittsburgh, Pa.

Lynch, Andrew F., Statisticial Division, A. B. Leach and Company, 57 William Street, New York, N. Y.

McChesney, Ruth, Agricultural Economic Research, University of California, Berkeley, Cal.

McDonald, W. Stewart, Analyst, Goldman, Sachs and Company, 30 Pine Street, New York, N. Y.

Miller, Ralph Gifford, Investment Banking, Clark Dodge and Company, 51 Wall Street, New York, N. Y.

Molleson, Dean C., Statistician, Munds and Winslow, 17 East 42 Street, New York, N. Y.

Popenoe, Herbert, Analyst, J. David Houser and Associates, 122 Fifth Avenue, New York, N. Y.

Post, Robert E., Assistant Professor of Economics, South Dakota State College, Brookings, S. Dak.

Rose, Dwight C., Investment Counsel, 10 Post Office Square, Boston, Mass.

Sanford, Horace L., Federal Reserve Bank, New York, N. Y.

Schur, George, Analyst, Publix Theatres Corporation, Paramount Building, New York, N. Y.

Shaw, Elbridge L., Dewey, Bacon and Company, Stocks and Bonds, 50 Broadway, New York, N. Y.

Stiles, Albert I., Statistician, Burnham, Herman and Company, 44 Wall Street, New York, N. Y.

Temnomeroff, Vladimir A., Assistant in Statistical Laboratory, Harvard Business School, Soldiers Field, Boston, Mass.

Urdahl, Dr. T. K., 5233 Locust Street, Philadelphia, Pa.

Viles, Charles L., Student, Harvard Business School, Soldiers Field, Boston, Mass.

Weber, Paul J. A., American Telephone and Telegraph Company, 195 Broadway, New York, N. Y.

Weller, M. Kingdon, First National Trust and Savings Bank, Los Angeles, Cal.

Wells, Carl D., Instructor in Social Statistics, Texas Christian University, Fort Worth, Texas.

Whitney, Simon N., Department of Economics, Yale University, New Haven, Conn.

Woods, Howard O., Market Forecasting, General Electric Company, Pittsfield, Mass.

Woolston, Dr. Howard B., Professor of Sociology, University of Washington, Seattle, Wash.

REVIEWS

Business Cycles: The Problem and Its Setting, by Wesley C. Mitchell. New York: National Bureau Economic Research. 1927. xxii, 489 pp.

On casual examination, Professor Mitchell's new book is likely to prove disappointing. We have awaited eagerly the publication of this volume, the fruit of twenty years' labor devoted largely to this one subject, in its various aspects, by one of the most brilliant of modern economists. Many of us have expected to find in it a comprehensive account of the play of interacting forces in each of the several stages of business cycles; a work which, for all except the few who have devoted nearly as much time to the problem and brought to it something like the same mental equipment as Professor Mitchell, should bring a revolution in understanding of business cycles. The title is disappointing, for it announces the work as a statement of the problem, where we had looked for a solution, not complete indeed, but as complete as existing knowledge, brought together by a master, could make it. For this solution, it appears, we must await a second volume on *The Rhythm of Business Activity*.

The table of contents bears out the title. The first chapter on "The Processes Involved in Business Cycles" is devoted chiefly to an account of the theories now current (pages 11–47), built on a background of earlier theories, and to plans for further work (pages 47–60). The second chapter is a summary of the characteristics of modern economic organization, as they influence business cycles. Chapter III, "The Contribution of Statistics," proves to be largely a summary of the statistical methods that have been used in studying business cycles (pages 193–269), followed by a presentation of eleven indexes of business conditions prepared by various workers, and a fifty-page summary, in two sections, setting forth the characteristics of business cycles as reflected in these indexes; but it contains little suggestion of analysis of the interplay of forces. Chapter IV reproduces, with but one significant addition,[1] the introduction which Professor Mitchell wrote for Dr. Thorp's *Business Annals*. Chapter V, with which the volume ends, is precisely what one should expect from the title and what has gone before, a statement of the problem in the form of "A Working Concept of Business Cycles" and a definition, followed by an outline of a plan for further work.

As one reads the first chapter, any disappointment which may have followed the first examination of the volume gives way to a feeling of admiration and satisfaction. This chapter proves to be neither a history of theories of business cycles nor a comprehensive summary of current theories, but a review of only so much of both as is necessary to determine, from the accumulated results of the labors of many economists over more than a century and a quarter, how to proceed in further advancing our knowledge. It brings the reader both to a hearty acceptance of the statement with which the work opens, that, "As knowledge of business cycles grows, more effort is required to master it," and to a conviction that

[1] A discussion of Dr. Frederick C. Mills' hypothesis regarding the relation of the length of business cycles to the stage of industrial organization (pp. 412–16).

the effort required, minimized by Professor Mitchell's careful selection of materials and lucid exposition, is well worth while.

Throughout the work, one who pauses to reflect must be impressed with the pains Professor Mitchell has taken to conserve the time and mental effort of the reader. Pertinent problems are not omitted and difficulties are not glossed over, but out of a vast fund of material, he picks out that which appears most important and useful for the purpose, organizes it carefully, and presents it in a manner to carry the reader steadily forward with sustained interest. The reader is never left long in doubt as to the significance of any portion of the discussion or of its place in the general scheme. The careful organization and rigorous suppression of details interesting for themselves, but of doubtful significance for the general picture, is in marked contrast to the ill-coördinated assembly of detailed discussions into which a writer is likely to be led when he has made a close analysis of a great mass of material. In this work one must read between the lines to perceive the erudition of the author.

Chapter I closes with a series of observations based upon or suggested by the study of theories of business cycles. Professor Mitchell concludes that "The conception of business cycles as a congeries of changes in numerous processes running abreast or following each other—a conception made vivid by the review of current theories—shows the need of quantitative knowledge. Indeed, our best chance of improving upon the work of earlier writers lies in this direction." History also, he believes, must be drawn upon heavily. These points are further discussed in the last section of Chapter III. The character of general business cycle studies must gradually change, Professor Mitchell holds, from attempts to explain business cycles in terms of cause and effect to analytic description. Finally, he considers the alternative, in proceeding with the study, of investigating the several theories one by one or of "using the theories to suggest facts which should be gathered and relations looked for," seeking primarily conclusions regarding the fluctuations and "treating verdicts upon the theories as by-products to be turned out when convenient." The latter plan he chooses without hesitation.

In Chapter II Professor Mitchell considers first the historical development of business cycle fluctuations to determine the characteristics of modern economic organization which created conditions favorable to the emergence and continuance of business cycles. He concludes that business cycles are only incidentally conditioned on a capitalistic organization; that the essential conditions are a "money economy" or, as he prefers to call it, a "business economy." The historical review is most interesting.

The next hundred pages are devoted chiefly to an outline of the modern business economy: the degree of concentration on production for profits; prices and their interrelations; the circulating medium and its characteristics; the flow of money payments; the elusive "entrepreneur"; and international differences in economic organization. Even Professor Mitchell's genius as a writer is not equal to the task of making this hundred-page summary of modern business organization absorbingly interesting to an economic theorist. As a descriptive summary, it is extraordinarily complete, and there is more of theoretical signif-

icance in it than one might expect. A few of the high lights are worthy of notice. Among economic resources, first place is given in the discussion to the knowledge and skill of the population. Granting that the value of this knowledge and skill cannot be appraised, Professor Mitchell regards its cost of maintenance as an important item which he roughly estimates. The subject of the flow of money incomes is treated more broadly than is usual, and with an important shift in emphasis arising from the demonstration that the flow of payments among business enterprises is several times as large as the flow of incomes to consumers and back to business enterprises, and varies much more through the successive stages of business cycles. The "entrepreneur" in modern large scale business is hunted down and revealed to view more adequately and convincingly than in any other economic study I recall. It appears, of course, that the "entrepreneurial function" has in its turn succumbed to the persistent pressure toward division of labor. Chapter II closes with a two-fold summary and a plea for a revised conception of economic equilibrium. The analogy with mechanical equilibria should be abandoned, it is held, and in its place should be substituted the concept of a balance-sheet equilibrium.

The summary of the development of statistical studies and of the statistical methods commonly used, occupying the first eighty pages of Chapter III, seems to me truly extraordinary. The declared purpose of the chapter is "to show what problems the statistical workers have attacked, how far they have pushed their researches, and what light they shed upon the character of business cycles."[1] Professor Mitchell specifically disclaims competence to give a full exposition of the statistical methods themselves. Nevertheless, the review of the principal statistical methods employed in business cycle research, their uses and limitations, is so thorough that I believe no economic statistician will fail to find in it stimulating suggestions. As an introduction to the statistical analysis of time series, covering the question of the place of statistics in economic research and its development, the adequacy of existing statistics, and the merits and defects of the various methods of measuring secular trends, seasonal variations, and cyclical and irregular fluctuations, I have not seen its equal. In addition to providing a more comprehensive review of these subjects than has been brought together anywhere else, to my knowledge, it includes a number of original contributions. The emphasis throughout is upon the ideas involved in the statistical methods and upon their adequacy for the problems attacked. No equations or mathematical formulae are to be found,[2] but the substance of the methods is set forth more completely than in any technical treatment I have seen.

This section of the work constitutes a most interesting experiment along lines which a few economic statisticians, at least, have thought should be followed quite generally: the presentation of an introduction to statistical methods in the form of a discussion of ideas and concepts with very limited treatment or, as in this case, no treatment at all of the details of statistical technique. The usual objection to such a proposal is that the ideas and concepts cannot be grasped without a knowledge of the details of the methods and an acquaintance with the mathe-

[1] P. 202.

[2] Exception must be made of two or three footnotes and equations of lines of trend shown on charts.

matical derivations and proofs on which they rest. Professor Mitchell has here presented more of the ideas and concepts involved in the statistical methods he discusses than are included in any of the statistical texts and with no consideration whatever of the details of statistical technique. Whether the presentation is effective for those not already versed in the technique, no statistician is perhaps competent to say on his own authority. The first step has been taken in an important experiment; the next step is to learn the reactions of readers who have not been trained in statistics.

In the whole discussion of statistical methods the reviewer finds nothing to which he is disposed to take exception, save one point in the brief discussion of correlation. Professor Mitchell contends that "coefficients of correlation are not percentages, though the fact that they run on a scale from -1.00 to $+1.00$ has seemed to Professor Fisher sufficient warrant for calling them percentages,"[1] and in a footnote remarks that, "If we insist upon using a percentage scale it should be that of the squared coefficients."

Here, as in so many cases to which Professor Mitchell calls attention, there is more of truth in a combination of two views than in either one separately. Professor Fisher is entirely correct in his statement. The coefficient of correlation does in fact give, with mathematical accuracy, the ratio of the variation in one variable which may be "explained" by the variations in another variable, to the total variation in the one variable (as variation is commonly measured, in terms of standard deviations). But so also is Professor Mitchell correct in his statement, extended in a footnote, that, "In the perfect case of exhaustive explanation" by several strictly independent factors "it is not the several coefficients which equal 1.00 but the sum of their squares."

We may agree, in general, with Professor Mitchell that "In considering how much significance attaches to a given coefficient of correlation . . . one should take the square of that coefficient rather than the coefficient itself," but it is important also to recognize the existence of the anomaly illustrated by the fact that both Professor Fisher and Professor Mitchell are technically correct in these statements. The anomaly disappears if variation be measured by the square of the standard deviation (a measure coming into some use under the name of "variance"), but so long as variation is commonly measured by the standard deviation (or by the mean or the quartile deviation—measures which show closely similar, though not mathematically identical, relations in this respect) the anomaly deserves to be emphasized.

Following the discussion of the technical problems attacked by specific statistical methods, Professor Mitchell proceeds to an examination of eleven indexes of business conditions and an illuminating discussion of differences in their movements. The differences among indexes for the United States he attributes chiefly to differences in the "economic processes to which they relate," a variety in character which he regards as highly desirable. Valuing the indexes which relate to specific economic processes, the changes in which may be visualized concretely, Professor Mitchell sees also much use for the "general business index," despite the nebulous character of the concept involved.

[1] This and the following quotations are from p. 268 ff.

While Professor Mitchell's chief contribution toward an understanding of business cycles in this work lies in his selection, organization, and interpretation of the great mass of theoretical and factual material elsewhere available, on two related points he presents an interesting analysis of his own. Having presented in Chapter III eleven indexes of business cycles, he proceeds to an analysis of the amplitude and the duration of the cycles in five of these indexes covering the longest periods of years. In Chapter IV he applies a similar analysis to data drawn from Dr. Thorp's *Business Annals*. Perhaps the most significant conclusions are: That there is no evidence of periodicity in business cycles, of long cycles that are multiples of short ones, or of "any regularity in the sequence of cycles of different lengths"; that there are significant differences in average duration of cycles in different countries and significant trends in duration in individual countries; and that the duration and the amplitude of business cycles appear to be resultants of a multitude of factors largely independent of each other.

The material which Professor Mitchell has assembled for this study of durations and amplitudes of business cycles offers a rich field for further research.[1] If there are significant differences in the average duration of business cycles in different countries and significant trends in duration in individual countries, as Dr. Mills' investigations indicate, there is need for constructing frequency distributions of deviations from averages and from trends. When this is done, the new distributions may conceivably lead to conclusions differing widely from those indicated by the distributions of the raw data which Professor Mitchell uses.[2]

Another line of research upon these materials which appears to me to hold great promise, relates to the form of the frequency distributions. A great number of them are strikingly skewed with the mode well below the mean. Professor Mitchell calls attention to the fact that the frequency distribution of durations compiled from the combined data for all the business cycles covered in the *Annals*, conforms closely to a logarithmic frequency curve. If this conformity is proved to be significant, it should form the basis for some important conclusions, or at least for hypotheses well worthy of further study.[3] To illustrate: The present reviewer discovered some years ago that certain frequency distributions of costs of production of wheat, which are strongly skewed in their original form, become approximately normal when costs are expressed in terms of bushels of wheat produced per dollar, suggesting that the multitude of factors responsible for variations are much more nearly mutually independent in their effects on yield of wheat per dollar of outlay than in their effects on cost per bushel. An interesting illustration from another field is furnished by a study of the volumes of phytoplankton obtained in a series of collections from the sea off

[1] Professor Mitchell comments briefly on this fact; see especially p. 420, footnote 4.

[2] Until closely analyzed, the evidence of significant differences in averages and of significant trends is very questionable. It was therefore natural that Professor Mitchell should have disregarded them in his first analysis. I am not sure that Dr. Mills is even yet confident of his ability to determine averages and trends for individual countries with great accuracy.

[3] One of the earliest and, I believe, still one of the best studies of the theoretical significance of frequency distributions differing from the "normal" has passed almost unnoticed by economic statisticians. Anyone dealing with such a problem should read the paper on "Skew Frequency Curves in Biology and Statistics," by the distinguished Dutch astronomer, J. C. Kapteyn (Noordhoff, Groningen; Wm. Dawson and Sons, London, 1903).

the coast of southern California, made by Dr. Ellis L. Michael of the Scripps Institution for Biological Research.[1] The frequency distribution is J-shaped when based upon the original observations, but becomes approximately normal when based on the logarithms of the data. These facts suggest that the multitude of factors which affect the quantity of phytoplankton in a given volume of sea water operate quite independently upon the rate of multiplication of the phytoplankton, but (a necessary corollary) not at all independently upon the volume of phytoplankton.[2]

The form of Professor Mitchell's frequency distributions suggests to me, however, the possibility that the true explanation may be found, in large measure, in quite a different quarter—in a theory of business cycles which has received attention so recently that it is mentioned by Professor Mitchell only in a paragraph in the Addenda summarizing a recent paper by E. E. Slutsky.[3] The form of several of the distributions is precisely what one should expect if business cycles result in considerable measure from a cumulation of the influence of mutually independent chance factors. We have too much evidence of the interaction of economic forces in business cycles to accept such a theory as a complete explanation, but there may be more truth than is now generally admitted in the "episodic" theory of business cycles, as thus revised.[4]

No review of Professor Mitchell's new book would be complete without special mention of the numerous observations on methods of research, usually with special reference to economic research, which are scattered through the text and in footnotes. For the professional economist they represent a significant and interesting feature of the work.

HOLBROOK WORKING

Food Research Institute, Stanford University

The Iron Industry in Prosperity and Depression, by Homer B. Vanderblue and William L. Crum. Chicago: A. W. Shaw Company. 1927. 193 pp.

The iron industry has long been regarded as a significant barometer of trade. It has even been shown that a single steel plant is a sensitive indicator of general business conditions.[5] There can be no questioning its strategic importance,[6] though perhaps some of our high regard for it as a barometer may have a prag-

[1] "Effect of Upwelling Water upon the Organic Fertility of the Sea in the Region of Southern California," *Special Publications of Bernice P. Bishop Museum*, No. 7, 1921. See especially pp. 555–66.

[2] I take the responsibility for this interpretation, which represents one obvious hypothesis suggested by the facts. Dr. Michael ventures no hypothesis on this point in the paper cited, his interest centering on the evidence thus afforded that in the subsequent analysis of the data he should deal with the logarithms of the volumes.

[3] P. 478.

[4] A suggestion offered by Dr. Harold Hotelling in a recent paper on "Differential Equations Subject to Error, and Population Estimates" (this JOURNAL, September, 1927; see especially pp. 289–91) may prove useful in investigating one phase of this problem.

An interesting illustration of the cyclical appearance of a curve constructed by cumulating random movements is to be found in a paper by Professors Bullock, Persons, and Crum on "The Harvard Index of Business Conditions," in the *Review of Economic Statistics*, IX, 2, April 1927, p. 82.

[5] W. L. Crum, "An Iron Producer and Business Cycles," *Harvard Business Review*. Vol. 5, p. 298.

[6] Warren M. Persons, "The Iron and Steel Industry during Business Cycles," *Review of Economic Statistics*. Vol. 3, p. 378.

matic origin, in that it is one of the very few industries wherein long-time data are available. It is to this industry, with its excellent statistical material, that Professors Vanderblue and Crum have chosen to apply their experience in statistical analysis.

The data are monthly pig iron and steel production and prices from 1902 to date, for the country as a whole and for regions. About one-half the book is devoted to examination of the regional data. The last five chapters introduce other series as possible forecasters. The method of analyzing the series is that regularly used by the Harvard Economic Service. In general, the method does surprisingly well for such wholesale application—and where the straight line trend is unfortunately a poor fit, the authors manfully stand by their method.

> . . . in the instance of the Lehigh district . . . the straight line fails to give a good "fit." . . . A curved trend was not used in the actual analysis, partly because no reason appears to expect that further decline in the immediate future which the use of such a curve would imply, and partly because the use of the straight line facilitates the comparison with the records for other districts. (Page 54.)

In certain cases, where the post-war trend is utterly different from that before the War, the authors use a broken trend. Those who quarrel with the link-relative method of correcting for seasonal fluctuations, can waive this case, for the seasonals are slight. For the country as a whole, making allowance for the length of month, the extremes are 97 and 104. Incidentally, while discussing the method, it is interesting to note that, although the study is one of relationships and similarity between series, correlation coefficients are never used.

What do we learn from this process? At first glance, the book is awe-inspiring, with its over-size format, its many tables and charts, and its appendices. At second glance, it is disappointing—one must hunt to find the conclusions for there are no convenient summaries, and the authors seldom stray from the immediate data into the economic implications of their discoveries. They are avowedly interested in forecasting. At third glance, it is stimulating, a veritable mine of data and technical comment, most provocative to one with the inquisitive instinct. It raises more problems than it solves. For example, let us consider two rather surprising facts which the analysis uncovers.

First, it is generally known that, contrary to the price behavior of agricultural commodities, the price and production of pig iron, or steel, move up and down together. But close examination of the charts and of statements made by the authors, indicates that this is not exactly true. Whereas, on the average, the price of pig iron reaches its cyclical peak slightly before pig iron production, it does not turn up in time of revival until ten months after the up-turn of production. More elaborate study on this point by Dr. Frederick C. Mills, in his inquiry into the behavior of prices for the National Bureau of Economic Research, indicates that for cycles between 1890 and 1925, on the average, pig iron prices rise 14.8 months and fall 23.0 months, whereas pig iron production rises 28.4 months and falls but 11.5. The difference between the cyclical behavior of production and prices is certainly striking and suggests a pretty problem which may not be limited to the iron industry alone. It would appear, in the case of pig iron, as though the price decline begins approximately at the time of recession

in production, and continues not only through the period of lessening production, but until production has recovered notably. For six cases where data were available, it appeared that capacity had advanced from an average low of 37 per cent active to approximately 50 per cent active before price revival developed. In other words, prices do not advance until capacity employed increases so much that certain producers are able to emplcy price increases as a method of selection among prospective purchasers without restricting their sales below their own capacity.

A second interesting fact made clear in the study is that there has been a marked reduction in the amplitude of pig iron price fluctuations. This change, specially apparent about 1907, has already been noted by Blackett.[1] According to Berglund [2] even wider price fluctuations occurred in 1897–1900, so the price stabilization process has been going on for some time. What is the explanation for this tendency? The authors have two suggestions, both derived from Blackett:

1. The growth of pig iron capacity between 1898 and 1907 did not keep pace with the rapid increase of steel furnace capacity, resulting in enhanced competition among pig iron purchasers, and runaway markets of wide amplitude. But a changed rate of expansion in the two industries after 1907 corrected the maladjustment.

2. The share of the output contributed by furnaces controlled by steel works has increased steadily, resulting in better coördination of producing programs of blast furnaces and steel works, and smaller accumulation of inventories.

3. A third suggestion made by Blackett has been disregarded, possibly because recent experience appears to discredit it—that the reduction in amplitude was due to lowering the tariff, inasmuch as the reduced duties and consequent enlarging of the area of effective market were coincident with the disappearance of runaway markets.

4. A slight technical influence may be suggested—that since the trend for pig iron prices is downward, as the authors point out in another connection, and since, in general, prices appear to fluctuate geometrically rather than arithmetically, this decreased amplitude may be in part, at least, a matter of definition.

5. An important factor is uncovered when production is examined by regions, for it becomes evident that certain regions are dominated by old, high-cost furnaces, described by the authors as "marginal." These areas have much wider cyclical fluctuations than those where furnaces are new and low-cost. Furthermore, the fact that they lag in revival indicates that they are high up on the supply schedule. But the importance of these high-cost furnaces has steadily decreased during the period. The replacement by low-cost furnaces of much larger capacity has resulted in a less scattered supply schedule, and a more stable price. The same condition appears in the coke industry, where beehive coke represents the old high-cost process, and by-product the new. Similarly there, the fluctua-

[1] O. W. Blackett, "Some Price Determining Factors in the Iron Industry," *Review of Economic Statistics*. Vol. 7, p. 203.

[2] Abraham Berglund, "The United States Steel Corporation and Price Stabilization," *The Quarterly Journal of Economics*. Vol. 38, p. 1.

tions are much wider for beehive than by-product coke. The implication in these data is that there is a tendency for price stabilization present in uniformity of costs, such as develops in large-scale industries where the technical process is public property and relatively stable.

6. Finally, an industry may meet the incidence of the business cycle by either price variation and relatively stable production, or price stability and wide fluctuations in production. In general, business men have adopted the first method, endeavoring by price adjustments to maintain, or revive production. Apparently, the second is the procedure in the iron industry. In fact, the policy of price stabilization was openly avowed by the United States Steel Corporation in 1921–22. An examination of the data in *The Iron Industry* shows that the reduced amplitude in pig iron prices finds no corresponding reduction in the fluctuations of production, which are wider, if anything. Never before was there a decline in production as marked as that of 1921, and the reduction in output in the minor recession of 1924 to less than one-half that of months in the previous year, is approximately that of such earlier depressions as 1908 and 1894. The tendency is towards production cycles rather than price cycles. From the point of view of the iron industry, it probably tends to prevent speculation, and thus create some degree of actual stabilization. But, from the point of view of those theorists who advocate higher interest rates or higher wages during expansion as a means of preventing over-development, the undue expansion of other industries will be aided by stable costs of such a general raw material as iron and steel. Nor will the maintaining of high prices during depression aid in bringing revival to other industries. Furthermore, the transfer of the incidence of the cycle from prices to production, spells less regular employment for workers, although employment does not fluctuate as widely as production in the iron industry. Certainly, the implications of this tendency need examination.

These two examples are sufficient to indicate problems which the reviewer wishes the authors had attacked. But they are interested chiefly in forecasting, and doubtless felt under the necessity of curbing their interest in by-product problems. In that field, their results are quite definite. They find no series outside the industry which can be used for forecasting either iron and steel production or prices, though they only examine those which bear some logical relation to the industry. Therefore, the industry must look to itself. And there, although it differs somewhat in different periods and phases of the cycle, they find that net changes in unfilled orders precede scrap prices and United States Steel common stock, which in turn precede iron and steel prices, production, and net earnings of the United States Steel Corporation.

WILLARD L. THORP

Amherst College

The Workmen's Compensation Problem in New York State. New York: National Industrial Conference Board, Incorporated. 1927. xx, 375 pp.

"With full recognition of the inadequacy of the existing data for the purpose of arriving at final or conclusive judgments in regard to any of the more impor-

tant problems involved in workmen's compensation, an attempt has been made . . . to present a survey of the more significant features of the operation of the workmen's compensation legislation and administrative machinery in the leading industrial state, New York. . . . The study does not attempt a minute analysis of detail, which in so complex and technical a subject would obscure rather than clarify the picture; it offers, rather, a survey of the more important questions which New York State presents in this regard" (page 8). "The main purpose of this study was to compare the cost of workmen's compensation under the present New York compensation law with that under the laws of New York's chief competitor states" (page 322).

To these ends the National Industrial Conference Board presents *The Workmen's Compensation Problem in New York State*, a compilation of the readily available information on the development, present status and operation of the workmen's compensation law in that state. After explaining the present provisions of the law and their origins, and comparing the law with those of other states, the study singles out for special attention the attitude of the courts, workmen's compensation insurance and its problems of rate making and reserves, experience in the operation of the law, and the cost of the law to employers in New York compared with the cost in other states.

In Chapter III is presented a careful analysis of court decisions in appeals from the decisions of the Industrial Board and its predecessors. The conclusion is reached that "while the burden of responsibility borne by the employer has been increased, the administration of the law has not been unreasonable or unfair. But a large part of the credit for the equitable character of that administration is due to the restraint imposed by the judicial review of compensation cases" (page 113). It is stated that "the fact that no appeals are made in state fund cases undoubtedly tends to make the Industrial Board and its referees less cautious in making such awards. Moreover, it is probable that there is less investigation by the state fund to prevent fraudulent claims than in the case of private carriers" (page 45). Both of these statements may be true, but no definite evidence is presented in support of them. The opinion is advanced that "the acceptance of the theory that industry should bear the cost of compensating accidentally injured workers does not serve to justify awards of compensation in cases where injury is due solely and entirely to the fault of the injured employee" (pages 106–7). This opinion raises the question whether the authors of the report would favor the inclusion of a provision in the law relieving the employer of liability in such cases. There are involved not only considerations of abstract justice but the eminently practical consideration of whether such a provision would not lead in too many cases to raising the ghost of "personal fault" with its attendant evils.

The explanation of the various types of insurance carriers and the discussion of rate making and reserves are, in general, well balanced and adequate. Much more attention has been devoted, however, to the arguments against monopolistic state insurance than to those in its favor. Certain isolated points have raised questions in the mind of the reviewer. It is stated that "The term self-insurance is obviously a misnomer, because there is nothing in the system that can be con-

strued as serving the same end as insurance, which is primarily intended to distribute losses" (page 123). It is submitted that if an employer is in a position to maintain a self-insurance fund on a sound insurance basis and does so, the term self-insurance is correctly applied. Certain employers have created their own insurance carriers to write only their own risks. It seems to the reviewer that the prime function of insurance is the replacement of uncertainty by certainty. This may be accomplished within a given employer's organization, as well as by grouping together a large number of different employers. It is stated that "Stock companies can call upon their stockholders for additional working funds in case of need to the extent of their holdings, and it is only as a final measure that the policyholder assumes the liability" (page 170). This seems to imply that stock companies may enforce the collection of an assessment equal to the amount of stock owned. This is not a fact.

Chapter V, dealing with the operation of the law, is perhaps the most interesting. As a basis for this chapter the Conference Board sent out a large volume of questionnaires to employers, to employees' organizations, to groups in contact with labor and to insurance carriers. Administrative and technical bodies were freely consulted. Unfortunately no usable information was received from wage earners, or from those who might represent them, nor apparently was any attempt made to secure their point of view by other means. After a fairly complete statistical picture of the industrial accident situation, past and present, which includes much of an academic nature, the report passes to a consideration of administrative and medical problems. Numerous suggestions are made for improving the methods in vogue in these latter departments, most of them originating with employers or insurance companies. The authors find the referee system especially open to criticism. It is here that one misses the complete discussion which broader sources of information would have made possible. Accident prevention and insurance are briefly treated. The statistics showing trends in the attitude of employers toward various systems of insurance are particularly valuable. In discussing self-insurance it is pointed out that 48 companies had deposited as security $1,632,927, "which has been withdrawn . . . from productive uses" (page 235). While this amount cannot be used by the companies in question it is hardly correct to say that it has been withdrawn from productive uses since it is represented by securities. It is further to be noted that insurance companies would have been obliged to segregate funds for the same general purpose if self-insurance were not permitted.

The cost of workmen's compensation in New York is exhaustively set forth in Chapter VI, so exhaustively that he who runs is advised to read the summary on pages 326–328. But for one who is interested in a detailed comparison of available cost figures this chapter is to be highly recommended. It boils down to the statement in the succeeding chapter that "When considered from all points of view, the cost of compensation in New York is both relatively and absolutely higher than in any other large industrial state" (page 326). The authors of the report are particularly concerned with the effect of workmen's compensation as a factor in competition. They state that "From the replies of employers no definite conclusion can be reached as to the direct effect of compensation costs

upon competition with other states" (page 272), and that "while compensation costs in New York State are seldom sufficiently high actually to force an industry out of the state, they may be a considerable factor in dictating the advisability of locating new plants elsewhere" (page 274).

The reviewer is unable to agree with the following proposition: "If the employer is able, by means of capital expenditure or through better management, to reduce the number of wage earners actually required or to effect the same result by increasing production, he would seem to be entitled to the benefits which accrue in the form of lower direct and indirect costs per unit of product without having the profit margin reduced by a fixed, though in effect increasing, charge for compensation" (page 320). It is not suggested that the employer *should* have the profit margin so reduced, but rather that provision should be made for reasonable compensation regardless of the effect on profits. The well-founded conclusion is reached that it is "impossible to ascertain the exact facts relative to accident and compensation experience, and these facts are indispensable if the actual accident situation and trend are to be presented in their true proportions, and in a way which will make possible constructive remedial action" (page 321).

Taken as a whole this study is much to be commended. It brings together an immense amount of valuable information and presents it in systematic and understandable fashion. The reader is given an excellent picture of the problem in its principal phases with a notable freedom from bias in most instances. The reviewer has attempted to indicate those instances where he feels that this quality is not fully preserved.

The report serves further to bring out two important points: (1) That all comprehensive studies of workmen's compensation must present their information and conclusions considerably after the fact, and (2) that there is room and need for vast improvement in the gathering and compiling of statistics of industrial accidents and their compensation.

RALPH H. BLANCHARD

Columbia University

Food Costs and City Consumers, by Charles Enos Artman. New York: Columbia University Press. 1927. 170 pp.

This monograph deals with a subject of great interest—the large spread between the car-load cost of fruit and vegetables and the retail cost to the consumer. The area taken for investigation was the New York market, which includes besides greater New York a large part of Northern New Jersey. Within this great market, the author has selected some fifty scattered retail stores. For each of these, he has secured retail prices week by week for sixteen months for those vegetables and fruits most largely bought. During the same period, on similar dates, wholesale and jobbing prices were secured for the same articles. There are interesting descriptions of the organization of the market and the functions of the wholesalers, jobbers and retailers. The various kinds of retail stores are also described. But the chief interest of the book to the readers of this JOURNAL will be in the results of the study.

The spread was found to vary widely for different types of stores. The spread in unit stores (including jobbers' spread) is much greater, for example, than in chain stores. This difference is explained partly by the fact that the chain stores buy outside the market at less than the wholesale price and partly by the fact that in the chain stores perishables do not bear their proper part of the expense of distribution. But there remains a considerable amount attributable to the greater economy of the chain store.

The most important result of the study, however, is the author's finding as to the cause of the wide difference in the spread among the different articles. Northern potatoes, for example, show a spread between the wholesale and retail price of 37 per cent of the retail price, while white onions have a spread of 63 per cent. The author after trying out statistically various hypotheses finally reached the conclusion that the amount of the spread is determined by the size (in money) of the average sale of the article. The explanation of the wide difference in the spreads of potatoes and white onions, cited above, is therefore that customers buy more (in money) of potatoes at a time. The average sale of Northern potatoes is 26.7 cents; the average sale of onions is 20.3 cents. This hypothesis appears to be well supported by the tables given. The tendency of the retailer is to price his goods so as to secure the same money profit per sale.

The author's conclusions as to the possibilities of a reduction in the retail prices are, therefore, of two kinds. In the first place, he suggests that the unit stores by combination in buying and delivery may reduce considerably the spread. It seems more likely at present that the chain stores will gradually replace the unit stores. In the second place, since the spread is greater on articles sold in smaller (money) quantities, the suggestion is made that the consumers should buy in larger quantities. To bring this about, the author would have retailers adopt the plan of selling at definitely fixed scales of prices graded regressively according to the quantity sold. The result of this would be the elimination of the small stores and the reduction of the average spread. Undoubtedly this would be desirable. The only question is whether the consumer would avail himself of the opportunity in any large way. The consumer is a lazy fellow and, perhaps, always will be.

The monograph is an excellent piece of work. It is, as far as the reviewer knows, the first really competent performance in a field of great interest.

GEORGE E. BARNETT

The Johns Hopkins University

Twin Births and Twins from a Hereditary Point of View, by Gunnar Dahlberg. Stockholm: Bokförlags, A. B. Tidens Tryckeri. 1926. 296, xv pp.

With the coöperation of the Swedish State Institute of Race Biology this study has been made of twin births and twinning from the genetical standpoint. The first half of the work deals with the literature of the subject and discusses twin ratios and the relative occurrence of monozygotic and dizygotic twins among twin births. This last subject is a very difficult one in view of the more recent evidence that monozygotic twins are not always characterized by phaenotypic resemblance. The statistics in regard to twin ratio and age of parents and size of

family, legitimacy, density of population in the place of birth, racial differences in respect to twins and relative incidence of triplets and quadruplets are satisfactorily summarized. Some ten pages are devoted to twin theories and a new hypothesis of twinning is suggested. This hypothesis is that there is a doubling factor in the egg, or possibly the sperm, which is manifested at about the time of the reducing division. The second part of the paper deals with the genetical factor in twinning, including such traits as stature and weight of twins and hereditary dissimilarity in monozygotic twins. This, again, is largely a summary of the literature. The author has, however, made some observations on mirror-like resemblance of one-sided expression of bilateral characters in monozygotic twins. For example, functional superiority of one leg, in one of the twins, of the other leg in the other twin, similarly of the tendency to place the right thumb uppermost in clasping the hands. The author has measured a number of twins and his whole discussion of the differences between repeated measurements, and the evaluation of the results found is of great importance. The measurements made are printed in full in some twelve pages and the book concludes with an appendix of some eighty plates of photographs of twins, and especially of the ears of twins. The whole constitutes an important contribution to the study of twins, especially to the quantitative statistical study.

CHARLES B. DAVENPORT

Cold Spring Harbor, Long Island

Economic Status and Health, A Review and Study of the Relevant Morbidity and Mortality Data, by Selwyn D. Collins. United States Public Health Service Bulletin 165, September, 1926. Washington: Government Printing Office. 1927. 74 pp.

The author has put every student of social pathology under very great obligations for having brought together in this compact form a wealth of information as to relations of economic status and health, particularly as to "poverty" as a factor in determining sickness and mortality rates. The excellent bibliography covering some seventy titles indicates a very careful search of available material. That material is not any too comprehensive. It represents partly general data on vital statistics analyzed by economic status, and partly a number of special studies limited in scope but intensive in character. The significant thing which is brought out very clearly both in the general text and in the author's conclusions, is that practically without any exception every investigation into this problem corroborates the point of view generally held by all students that both morbidity and mortality are directly influenced by economic and occupational status, whether it be sickness in Hagerstown, Maryland, with a higher rate for the moderately well off than for the well-to-do, and a still higher rate for the poor, or a study limited to twenty-four South Carolina mill villages, showing an increase in the number of days of disabling sickness with a decrease of the weekly family income, or the workmen's sick and death benefit fund, or a comparison between taxpayers and non-taxpayers in Providence, Rhode Island. In every case, whether taken as a whole or distributed by age groups, the mortality and

morbidity show the effect of economic need. Quantitatively, the largest material is that dealing with mortality occurring to occupational groups in England and Wales, where professional and salaried people, skilled labor and unskilled labor, show an ascending mortality rate. Again, as would be expected, the same influences are even more manifest in dealing with child morbidity and mortality.

In discussing these very interesting relationships, the author inquires whether the "causative" association between poverty and death-rates is general or specific. He argues that the relationship must be considered as specific because it differs in intensity for different diseases. The influence of poverty appears to be very much stronger in influencing the death-rate from tuberculosis, bronchitis, pneumonia and accident, while in some diseases as, for instance Bright's Disease, no particular correlation can be found. And furthermore, there is a group of diseases in which the correlation between economic status and death-rate is positive, *i. e.*, the death-rate rises with the rise of economic status. That seems to be true of diabetes, gout, cirrhosis of the liver. Of course, the latter conclusion is not particularly startling in face of the general knowledge as to the causation of some of these diseases.

In passing, we may note the very curious "slip," which should require a careful explanation. Says the author (page 41), "If poverty is entirely the effect of sickness of the breadwinners of the family, we would expect a general relationship applying about equally to each cause of death, but such is not the case." The context leaves no doubt that what the author had in mind was the question whether poverty is entirely the *cause* of sickness and not the effect. Subconsciously, undoubtedly, the counter-proposition that poverty was not only the cause of sickness but also the result of it, was forcing itself to the surface of the author's mind.

The attention of eugenists should be called to the very interesting discussion as to how far heredity and environmental factors respectively may be held responsible for this seeming correlation between poverty and disease. It is, of course, very satisfying to some minds to insist that the higher death- and morbidity-rate among the poor is due to the poor being of an inferior stock. But as the author very properly points out, the inferiority of the stock then should manifest itself in an approximately equal degree in all the diseases rather than those which are obviously affected by conditions of living and by available medical aid. Why should the stock resist diabetes and gout better, but succumb to tuberculosis and bronchitis? If unskilled labor shows a mortality from accident three times as high as the professional class, it is obviously not due to the inferiority of the plasma. In approaching the same problem of the respective influences of heredity and environment in their influence upon infant mortality, the author properly argues that if the factor was hereditary, it should manifest itself from the very beginning of life. As a matter of fact, the difference due to economic status is very much stronger after the first month of life than it is during the first month. The presumption, therefore, is justified that the cause lies in the environment rather than the hereditary influences. The data for Baltimore, Maryland, are particularly interesting in this connection. Total infant mortality under one

month fluctuates according to economic status from 60 to 28, while for infants over one month but under one year, it fluctuates between 169 and 9.

The study closes with a somewhat more complicated analysis of the comparative influences of three factors—poverty, birth-rate and artificial feeding—based upon investigation of a certain number of rural districts in Bavaria. The statistical reasoning is somewhat too complicated to be recorded here in full. On a basis of partial correlation coefficients, it appears that birth-rate and artificial feeding exercise a stronger influence upon infant mortality than does general poverty. But as the author properly remarks, true as this deduction may be for the rural districts of Bavaria, it is doubtful whether the same would hold true of urban districts. For naturally, the difference in the economic status in the rural population does not exercise the same influence as in the city.

Altogether, the study is a most useful one in the present movement for improvement of public health conditions, medical practice, and in the revival of the movement for health insurance which may be noticed. It is, of course, very largely true that as one cynical social statistician has remarked, "Social statistics proves little except the obvious, and if it does, it is not to be trusted." But, after all, one must reckon with the fact that as a result of the processes of rationalization, even things obvious to the unprejudiced mind are frequently denied, and it is there that a clear statistical analysis may be helpful.

I. M. Rubinow

Jewish Welfare Society of Philadelphia

The Theory of International Prices: History, Criticism and Restatement, by James W. Angell. Cambridge: Harvard Economic Studies 28. 1926. xiv, 571 pp.

One of the most interesting tendencies in recent economic study has been the greater attention paid to doctrinal history. Tracing back, perhaps, to Böhm-Bawerk's influential *Kapital und Kapital-zins* this revival has attained a high plane of scholarly excellence in the work of Bonar and Cannan in England, of Halèvy and Weulersse in France, of Seligman and Taussig in the United States. A welcome consequence has been the attraction of a group of younger scholars to the study of *Dogmengeschichte*. The monographs of Turner, Furness and Munro are among the most creditable contributions which American scholarship has made to the progress of our science.

Into this pleasing company comes Dr. Angell with a laborious study of the development of the theory of international prices. Prepared as a doctoral dissertation, the essay was awarded the David A. Wells prize for 1924–25 and is now published from the income of that fund, under the direction of the department of economics of Harvard University.

The book is voluminous, extending with appendices and index, over some five hundred and seventy pages. Of this, roughly two-thirds is given, in approximately equal parts, to the development of English and of Continental thought. One hundred and ten pages are devoted to positive exposition. The appendices discuss recent studies of the English Bank Restriction, the international transfer of capital and European post-war experience.

The treatment is thus pitched in somewhat lavish key. The text in places suggests a transcript of diligently compiled notebooks, rather than a mature residuum. Francis A. Walker once complained that writing the "advanced course" of his *Political Economy* was easy enough, but that the "elementary course" was a real job—which, in turn, recalls Garrick's epigram: "Tragedy is simple; it's Comedy that's serious business."

In the historical section of Dr. Angell's monograph this diffuseness is most in evidence. Sometimes, as in the pre-Ricardian period, ground is retraversed that has been sufficiently surveyed. Elsewhere acute comment descends into meticulous hypercriticism, with the result that the reader is blinded with leaves and left groping for the forest.

The troublesome alternative which every historian of thought encounters between chronological sequence and topical grouping is met by Dr. Angell by parallel surveys of the two major areas, with some inevitable duplication, but on the whole with justifiable result. In the earlier period there is insufficient projection of doctrine against background, as against a disproportionate emphasis upon the disturbed experience of post-war trade. Secondary materials are used discriminately with the possible exception of wholesale reliance upon the studies of Silberling whose Heldian aberration as to Ricardo's intellectual honesty is deprecated with dove-like tenderness.

The positive sections, presented as "a restatement of the theory of international prices," recognize the main framework of the classical theory as valid, but make what are described as "serious inroads" and "great alterations" in its actual content. These modifications are (1) actual money costs and prices must be substituted for comparative labor costs as the basis of the doctrine; (2) regard must be had, in the doctrine of equilibrium in the mechanisms of international trade ("sound in its underlying principle") for changes in the demand and supply of bills of exchange as affected by the volume of bank deposits and the levels of general prices; (3) the classical doctrine of "the paper exchange" has only a very general and inexact application to conditions such as have been manifest since 1914, by reason of the effects of widespread speculation in the currency and the exchanges.

The doubt naturally arises as to whether Dr. Angell's painstaking labors, passing more than once into fine-drawn commentation, are not after all a skillful tilt against a phantom of his own making. The stoutest Ricardian will hardly claim that the classical doctrines of international value, any more than any other of the basic principles of our science, indeed of any science, prevail unmodified and unextended in periods and places widely different from those of their birth and nurture. Universalism and perpetualism are obsolete terms of attack that are properly assigned to the limbo of doctrinal dialectics. The real issue in the ascertainment of paternity and the appraisal of influence in the development of scientific thought is the quality which Alfred Marshall has defined so finely as "architectonic," and here Dr. Angell's incisive bite is less in evidence than his somewhat prolonged bark.

As to the justice of his strictures there is question thus, less of accuracy than of proportion. Professor Viner has broken a stout lance in defense of the indict-

ment of the classical doctrine and the last word has by no means been said. But these are matters that transcend the limits of a review. Dr. Angell has written a good book; with ruthless compression and more complete assimilation he would have given us a better one.

JACOB H. HOLLANDER

Johns Hopkins University

Banking Cycles, by Lincoln W. Hall. Philadelphia: University of Pennsylvania Press. 1927. 173 pp.

This book is described by the University of Pennsylvania Press in the following terms:

> The Next Step—Definite Forecasting
> Revolutionary—Sound—Inevitable
>
> . . . . .
>
> *Banking Cycles* marks the obsolescence of many of today's methods of measuring business cycles, and of all the present methods of forecasting by means of cycle analysis. It makes definite business forecasting a possibility, for the methods described for the computation of cycles and forecasts are applicable to all forms of commercial data in any type of business organization.

From these extraordinary claims the reader is led to expect an extraordinary volume, but instead he finds a very ordinary volume. The author attempts the familiar task of determining the cyclical variations of historical series, by eliminating secular trends and seasonal variations. The data in this case are the principal items of condition of national banks since 1878, and for more recent years of reporting member banks and the Federal Reserve Banks as well. The author does not attempt to bring to bear upon these banking statistics any related information, nor does he attempt to draw any significant economic conclusions from his analysis. He contents himself with various computations of cyclical variations; and his method of prediction is based upon the assumption that future cyclical fluctuations will repeat those of the past, without any modification due to changes in other related series.

Concerning the statistical methods which the author employs, it may be said that they are ingenious and are evidently the result of considerable labor. When, however, an author produces a book in a much-worked field such as this, it seems fair to expect from him some statement of his relationship to previous workers and a discussion of the points at which his methods differ from theirs, together with a presentation of some evidence to indicate that his methods are sounder than those generally accepted. Mr. Hall, however, in his book of 173 pages refers to only one other writer, and that incidentally, and presents no evidence to show that the methods he uses are superior to those used by others. It seems to the reviewers that the logical basis of his methods is decidedly unsatisfactory. For instance, the central principle of several of the methods is that both the ordinate of secular trend and the index of seasonal variation for any year can safely be based upon the figures for the three years immediately preceding. No computation upon so narrow a base will commend itself to those who have worked extensively on the problem of measuring secular trend and seasonal variation.

It would not be profitable to enter into a more detailed discussion of either methods or results. The book appears to the reviewers to contribute little to either statistical methodology or economic science.

There is no index, and the book is marred by a number of grammatical errors, as for example, in the opening sentence, "data . . . changes," and "phenomena . . . has."

ROBERT W. BURGESS
W. RANDOLPH BURGESS

New York City

Money Lending Practices of Building and Loan Associations in Ohio, by H. Morton Bodfish. Columbus: The Ohio State University Press. 1927. 84 pp.

This monograph gives a very clear picture of building and loan associations' lending practices in Ohio. The data were gathered by the questionnaire method, supplemented by interviews with associate officers. Of the total associations in Ohio, 36.3 per cent answered.

The subject is approached from the standpoint of pure practice. The statements are clear, and the work is well illustrated by charts. The astonishing hold such associations have on the investing public of Ohio may be gathered from the facts that in Dayton, the average per capita invested in building and loan associations amounts to $690.94, while Cleveland shows $168.64 and Cincinnati $241.45—an allowance must be made for the fact that Mr. Bodfish has used the 1920 census and the 1926 building and loan figures.

The distribution of mortgage loans by institutions shows that building and loan associations make up 47 per cent, while state, private, and national banks constitute only 33 per cent.

The maximum percentage of appraised value which associations will lend was found to be very close to the almost universal figure of 66⅔ per cent—80.4 per cent of the associations fall within this group, the median and modal figure being 66⅔ per cent. The importance of the per cent of appraised value that associations will lend is found to be considerably reduced by the wide variation in appraising. The questionnaire listed a hypothetical house with a fixed land value. Although the house was simple, and in low figure valuation, the appraisals of the association secretaries varied from $4,000 to $8,500, disregarding six extreme items. The average appraisal for the sixteen cities used, amounted to $6,600, while the average appraisal for the whole state is $6,143.

An attempt was made to find a correlation between the building and loan associations' assets per capita of population and the amount associations would lend on the hypothetical house; the results indicated that there is no such relation. Therefore, increased available funds do not appear to push the amount of individual loans far beyond general practice. In the larger cities both the appraised value of the hypothetical house and the maximum loans that would be granted, were considerably higher than in the state in general.

Loan costs, including legal fees, search, and appraisal, appear to be governed by habit or tradition, and fluctuate very little from year to year. Local variations

are, however, very great: Loan costs per $1,000 loan range from $6.27 in Sidney to $40.33 in Akron.

The building and loan assets per person did not appreciably affect the amount of the loan. The cities with the highest per capita assets do not consistently make the largest loans, but in many cases higher per capita assets are accompanied by low fees and vice versa. It would have been interesting if the study had included a detailed analysis to discover whether the low per capita assets caused the high charge per $1,000 loan; or whether the fact of high costs per $1,000 loan indicated a poorer community, or little interest in building and loan associations. It is natural to assume that an increase in loan funds would lessen the loan charges, and this factor is fairly well borne out by the figures tabulated.

The work of Mr. Bodfish should prove of great interest to building and loan associations in general, as it provides them with a careful study of lending practices, methods of appraisal and fees charged. It would be interesting to round out this study with an account of the borrowing methods and practices of obtaining funds; especially in comparison with the state savings banks.

JOHN G. ROLPH

The Amalgamated Bank of New York

Public Utility Finance, by Walter E. Lagerquist. Chicago: A. W. Shaw and Company. 1927. 671 pp.

This volume on *Public Utility Finance*, by Walter E. Lagerquist, is the first of a series of volumes entitled "Materials for Studies of Public Utilities," published by the Institute for Research in Land Economics and Public Utilities. The plan of the Institute contemplates other volumes on public utility economics, regulation, labor, accounting, operation and rate making.

Dr. Lagerquist in his comprehensive book has gathered material published in engineering and trade journals and in the proceedings of technical societies during the past fifteen years. In some cases the complete article is printed and in other instances only excerpts. These selections are preceded by chapter introductions and are linked together by comment and analyses by Dr. Lagerquist.

The subject matter covers a broad range and is grouped in six main divisions, *i. e.*, organization of public utility companies; general characteristics of utility securities and the fundamentals of the mortgage; capitalization structure, the regulation of public utility capital, the market for public utility securities and customers ownership; plan and investigation required in new utility financing; problems of operation including centralization of power supply, diversity factor, costs and depreciation charge; relation of public utilities to the business cycle, budgets and taxation problems.

In the selection of material Dr. Lagerquist has given considerable space to articles by writers who have had practical experience in dealing with the problems discussed. In the handling of controversial subjects the author has included articles representing diverse points of view.

The volume is a valuable source book for those interested in this important

subject. A bibliography should have been included. The work, however, was a very large undertaking and the selection has been carefully done.

H. K. HERWITZ

The Amalgamated Bank of New York

A Health Study of Ten Thousand Male Industrial Workers, by Rollo H. Britten and L. R. Thompson. United States Public Health Service Bulletin 162. Washington: Government Printing Office. 1926. v, 170 pp.

The purpose of this Bulletin is the coördination of the results of relatively complete physical examinations of approximately 10,000 male workers made by medical officers of the service incidental to surveys of a number of specific industries. The difficulties of this coördination were increased by the fact that no concrete program was mapped at the start for procuring comparable results so that naturally the survey of each industry was conducted with reference specially to the known hazards of that industry. The examining staff was also different for each survey—although there was a certain amount of overlapping.

Those examined included 1,453 workers in the pottery industry, 985 post-office employees in New York and Chicago, 1,511 workers in the glass industry, 353 employees of a gas making company, 1,317 foundry workers, 935 in the steel industry, 916 employees of a chemical company, 223 cement workers, 398 male workers in cigar factories and 2,052 male garment workers.

A preliminary discussion of the racial distribution in the industries surveyed is followed by a tabulation of the physiological measurements including blood pressure, pulse rate, vital capacity and chest measurements, height and weight. Next there is a discussion of the diseases and defects found. Finally we have a summary by industry and a final chapter giving general summary and conclusions. Detailed figures are given in twenty-five tables in the appendix.

ROBERT HENDERSON

Equitable Life Assurance Society

are, however, very great: Loan costs per $1,000 loan range from $6.27 in Sidney to $40.33 in Akron.

The building and loan assets per person did not appreciably affect the amount of the loan. The cities with the highest per capita assets do not consistently make the largest loans, but in many cases higher per capita assets are accompanied by low fees and vice versa. It would have been interesting if the study had included a detailed analysis to discover whether the low per capita assets caused the high charge per $1,000 loan; or whether the fact of high costs per $1,000 loan indicated a poorer community, or little interest in building and loan associations. It is natural to assume that an increase in loan funds would lessen the loan charges, and this factor is fairly well borne out by the figures tabulated.

The work of Mr. Bodfish should prove of great interest to building and loan associations in general, as it provides them with a careful study of lending practices, methods of appraisal and fees charged. It would be interesting to round out this study with an account of the borrowing methods and practices of obtaining funds; especially in comparison with the state savings banks.

JOHN G. ROLPH

The Amalgamated Bank of New York

Public Utility Finance, by Walter E. Lagerquist. Chicago: A. W. Shaw and Company. 1927. 671 pp.

This volume on *Public Utility Finance*, by Walter E. Lagerquist, is the first of a series of volumes entitled "Materials for Studies of Public Utilities," published by the Institute for Research in Land Economics and Public Utilities. The plan of the Institute contemplates other volumes on public utility economics, regulation, labor, accounting, operation and rate making.

Dr. Lagerquist in his comprehensive book has gathered material published in engineering and trade journals and in the proceedings of technical societies during the past fifteen years. In some cases the complete article is printed and in other instances only excerpts. These selections are preceded by chapter introductions and are linked together by comment and analyses by Dr. Lagerquist.

The subject matter covers a broad range and is grouped in six main divisions, *i. e.*, organization of public utility companies; general characteristics of utility securities and the fundamentals of the mortgage; capitalization structure, the regulation of public utility capital, the market for public utility securities and customers ownership; plan and investigation required in new utility financing; problems of operation including centralization of power supply, diversity factor, costs and depreciation charge; relation of public utilities to the business cycle, budgets and taxation problems.

In the selection of material Dr. Lagerquist has given considerable space to articles by writers who have had practical experience in dealing with the problems discussed. In the handling of controversial subjects the author has included articles representing diverse points of view.

The volume is a valuable source book for those interested in this important

subject. A bibliography should have been included. The work, however, was a very large undertaking and the selection has been carefully done.

H. K. HERWITZ

The Amalgamated Bank of New York

A Health Study of Ten Thousand Male Industrial Workers, by Rollo H. Britten and L. R. Thompson. United States Public Health Service Bulletin 162. Washington: Government Printing Office. 1926. v, 170 pp.

The purpose of this Bulletin is the coördination of the results of relatively complete physical examinations of approximately 10,000 male workers made by medical officers of the service incidental to surveys of a number of specific industries. The difficulties of this coördination were increased by the fact that no concrete program was mapped at the start for procuring comparable results so that naturally the survey of each industry was conducted with reference specially to the known hazards of that industry. The examining staff was also different for each survey—although there was a certain amount of overlapping.

Those examined included 1,453 workers in the pottery industry, 985 post-office employees in New York and Chicago, 1,511 workers in the glass industry, 353 employees of a gas making company, 1,317 foundry workers, 935 in the steel industry, 916 employees of a chemical company, 223 cement workers, 398 male workers in cigar factories and 2,052 male garment workers.

A preliminary discussion of the racial distribution in the industries surveyed is followed by a tabulation of the physiological measurements including blood pressure, pulse rate, vital capacity and chest measurements, height and weight. Next there is a discussion of the diseases and defects found. Finally we have a summary by industry and a final chapter giving general summary and conclusions. Detailed figures are given in twenty-five tables in the appendix.

ROBERT HENDERSON

Equitable Life Assurance Society

RECENT LITERATURE

STATISTICAL BOOKS

ASHBY, FORREST BEE. *The Economic Effect of Blue Sky Laws.* Philadelphia: University of Pennsylvania Thesis. 1926. 59 pp.

Bengal Chamber of Commerce for the Year 1926, Report of the Committee of the. Appendices. 1927. Calcutta: Criterion Printing Works. vi, 457 pp.

W. RANDOLPH BURGESS. *The Reserve Banks and the Money Market.* New York: Harper and Brothers. 1927. xxi, 328 pp.

ELDERTON, W. PALIN. *Frequency Curves and Correlation.* 2nd Edition. London: Charles and Edwin Layton. 1927. 239 pp.

FOOD RESEARCH INSTITUTE. *The World Wheat Situation 1926–27.* Wheat Studies. Vol. IV, No. 1. November, 1927. 62 pp.

——. *Statistics of American Wheat Milling and Flour Disposition since 1879.* Vol. IV, No. 2. December, 1927. 39 pp.

——. *Survey of the Wheat Situation August to November, 1927.* Vol. IV, No. 3. January, 1928. 31 pp.

FREDERICK, JOHN HUTCHINSON. *Federal Regulation of Railway Securities under the Transportation Act of 1920.* Philadelphia: University of Pennsylvania Thesis. 1927. 120 pp.

GARES, ROY L. *Immigration Restriction.* New York: The Macmillan Company. 1927. xi, 376 pp.

HANAU, ARTHUR. *Die Prognose der Schweinepreise.* Vierteljahrshefte zur Konjunkturforshung, Herausgegeben vom Institut für Konjunkturforschung. Berlin. 1927. 41 pp.

HANSEN, ALVIN HARVEY. *Business-Cycle Theory.* Boston: Ginn and Company. 1927. x, 218 pp.

HARDY, CHARLES O. and COX, GARFIELD V. *Forecasting Business Conditions.* New York: The Macmillan Company. 1927. 434 pp.

JACK, D. T. *The Restoration of European Currencies.* London: P. S. King and Son. 1927. vi, 218 pp.

LEIGH, ROBERT D. *Federal Health Administration in the United States.* Harper's Public Health Series. New York: Harper and Brothers. 1927. 687 pp.

LEWIS, CLEONA. *The International Accounts.* Institute of Economics Series. New York: The Macmillan Company. 1927. x, 170 pp.

LOWRIE, JOSEPHINE A. *Merchandise Repossessions in the Instalment Furniture Trade.* Special Bulletin of the Bureau of Business Research, Ohio State University. November, 1927. 41, ii pp. (mimeographed)

MALMQUIST, K. G. *Investigations on the Stars in High Galactic Latitudes.* Meddelande Fran Lunds Astronomiska Observatorium. Ser. II. Nr. 37. 1927. 103 pp.

NATIONAL INDUSTRIAL CONFERENCE BOARD, INC. *Industrial Progress and Regulatory Legislation in New York State.* New York. 1927. xiii, 148 pp.

NEWSHOLME, SIR ARTHUR. *Health Problems in Organized Society.* London: P. S. King and Son. 1927. xi, 253 pp.

NYEHARA, S. *The Industry and Trade of Japan.* London: P. S. King and Son. 1926. xv, 326 pp.

PEARL, RAYMOND. *The Rate of Living.* New York: Alfred A. Knopf. 1928. 185 pp.

RIETZ, H. L. (Editor-in-Chief). *Handbook of Mathematical Statistics.* Translated in Russian by S. P. Bobroff. Moscow. 447 pp.

RITTER, VON KURT. *Qualitätskontrolle und Standardisierung Landwirtschaftlicher Produkte.* Agrarpolitische Aufsätze und Vorträge. Berlin: Paul Parey. 1927. 64 pp.

ROBINSON, DANIEL SOMMER. *Illustrations of the Methods of Reasoning.* New York: D. Appleton and Company. 1927. 346 pp.

ROUMANI, ADIB. *Essai Historique et Technique sur la Dette Publique Ottomane.* Paris: Marcel Giard. 1927. 332 pp.

SELEKMAN, BEN M. *Postponing Strikes.* New York: Russell Sage Foundation. 1927. 405 pp.

SELIGMAN, EDWIN R. A. *The Economics of Instalment Selling.* New York: Harper and Brothers. 1927. Vol. I, xii, 357 pp. Vol. II, 623 pp.

STATISTISCHE MEDEDEELINGEN VAN HET BUREAU VAN STATISTIEK DER GEMEENTE AMSTERDAM. *Huishoudrekeningen van 212 Gezinnen uit Verschillende Kringen der Bevolking.* Amsterdam: J. M. Meulenhoff. 1927. x, 153 pp.

WALLACE, WILLIAM. *Business Forecasting and Its Practical Application.* London: Isaac Pitman and Sons. 1927. xiv, 117 pp.

WOODWARD, W. H. *Profits in Bank Stocks.* New York: The Macmillan Company. 1927. 112 pp.

GOVERNMENT PUBLICATIONS

UNITED STATES.

—— Bureau of Labor Statistics.

No. 445. "Retail prices, 1890 to 1926."

No. 450. "Wages and hours of labor in the boot and shoe industry, 1907 to 1926."

No. 451. "Safety code for forging and hot metal stamping."

No. 452. "Wages and hours of labor in the hosiery and underwear industries, 1907 to 1926."

—— Treasury Department.

Public Health Reports:

Vol. 42.

No. 42. "Studies on Efficiency of Water-Purification Processes"; "Co-operative Rural Health Work in 1926–27."

No. 43. "Amendments to International Rules of Zoological Nomenclature."

No. 44. "Report on Pellagra in the Mississippi Flood Area."

No. 46. "Report on Survey of Endemic Goiter in Oregon."

No. 48. "Life Expectancy in the United States and in England."

FOREIGN COUNTRIES.

India.

"Statistical Abstract for British India, 1916–17 to 1925–26." Calcutta: Central Publication Branch. 1927. xi, 707 pp.

Java.

"Prices, Price Indexes and Exchange Rates in Java, 1913–1926." Departement van Landbouw, Nijverheid en Handel. Batavia: Drukkerijen Ruygrok and Company. 1927. xxxii, 173 pp.

ARTICLES IN PERIODICALS

MONTHLY LABOR REVIEW.

October, 1927. "Opposition of organized labor to tipping system"; "Results of arbitration proceedings under railroad labor act of 1926"; "Labor legislation in Uruguay"; "Accident experience in iron and steel industry"; "Building permits in representative cities"; "Entrance wage rates for common labor"; "Wages and hours in cotton-goods manufacturing."

November, 1927. "Increase in union wage rates"; "International Federation of Trade-Unions"; "Labor productivity in copper mining"; "Problem of the unskilled laborer with a large family"; "Accident experience of various industries in 1925 and 1926"; "Construction cost of dwellings in various cities."

December, 1927. "The work of the U. S. Bureau of Labor Statistics"; "Public service retirement systems: State employees"; "Explanation of the new index of wholesale prices"; "Increased labor productivity of a large steel plant"; "Child endowment in Australia"; "International comparison of real wages as of July 1, 1927."

THE REVIEW OF ECONOMIC STATISTICS, Harvard Economic Service.

October, 1927. "Review of the Third Quarter of the Year"; A. Aftalion: "The Theory of Economic Cycles Based on the Capitalistic Technique of Production"; W. B. Smith: "Wholesale Commodity Prices in the United States, 1795–1824"; M. K. Zinn: "A General Theory of the Correlation of Time Series of Statistics."

JOURNAL OF POLITICAL ECONOMY.

October, 1927. L. C. Marshall: "Offerings in Economics in 1925–26"; J. L. Laughlin: "Indian Currency since the World War"; C. F. Roos: "A Dynamical Theory of Economics"; C. E. Warne and M. E. Gaddis: "Investigation of Industrial Disputes"; H. A. Logan: "The Federation of Catholic Workers of Canada"; H. Schultz: "Mathematical and Statistical Economics."

December, 1927. G. G. Tunell: "Rate Making and Excess Income"; P. T. Homan: "The Impasse in Economic Theory"; H. A. Logan: "Federation of Catholic Workers of Canada, II"; F. B. Ashby: "Individual Cycles of Stock Prices"; J. E. Cummings: "Need for Co-ordination of Transportation."

POLITICAL SCIENCE QUARTERLY.

December, 1927. R. Moley: "Some Tendencies in Criminal Law Administration"; J. Dickinson: "A Working Theory of Sovereignty, I"; J. Ford: "The Enforcement of Housing Legislation"; H. S. Quigley: "Federalism and Foreign Relations in China"; F. S. Dunn: "International Legislation"; L. A. Weissberger: "Machiavelli and Tudor England."

THE ANNALS.

November, 1927. "Europe in 1927. An Economic Survey." By M. Patterson.

JOURNAL OF ACCOUNTANCY.

November, 1927. D. P. Musselman: "Instalment Accounts Receivable"; W. Mucklow: "Land-value Insurance"; J. H. Jackson: "Present Tendencies in Business Education."

December, 1927. "American Institute of Accountants Examinations, November, 1927"; F. H. Hurdman: "Some Phases of the No-par-value-stock Problem"; J. Lynde-Lockwood: "Public Accounting in Mexico"; G. H. Newlove: "Depreciation."

January, 1928. W. B. Campbell: "Coöperation between Accountants and Bankers"; L. R. Robinson: "Some Accounting Problems of Investment Trusts."

INTERNATIONAL LABOUR REVIEW.

October, 1927. G. A. Johnston: "Social Work and Labour Legislation"; W. Lauber: "Fifty Years of Legislation on Occupational Diseases in Switzerland"; M. Besson: "Labour Legislation in the French Colonies"; I. F. Ayusawa: "The Population Problem and Industrialization in Japan."

November, 1927. H. Fuss: "Money and Unemployment"; A. Stocker: "The Medical Aspect of Vocational Guidance"; H. Goldschmidt: "The Application in Germany of the Washington Convention concerning the Employment of Women before and after Childbirth"; E. Haultain: "Mothers' Allowances in North America"; D. Nikoloff: "The Trade Union Movement in Bulgaria."

December, 1927. J. Morellet: "At What Moment do the International Labour Conventions become Applicable?" A. Tixner: "Sickness Insurance at the International Labour Conference"; F. Rager: "Provision for Prolonged Unemployment in Certain Industrial States"; A. Stocker: "The Medical Aspect of Vocational Guidance: II."

JOURNAL OF THE ROYAL STATISTICAL SOCIETY.

Part IV, 1927. E. T. Elbourne: "Trade Association Statistics"; D. C. Jones: "Pre-War and Post-War Taxation"; L. J. Reed and R. Pearl: "On the Summation of Logistic Curves"; R. A. Lehfeldt: "The World's Stock of Money"; D. Frisch: "Sur un Problème d'Économie Pure."

THE RECORD. American Institute of Actuaries.

June, 1927. "Address by the President, J. G. Parker"; H. R. Corbett: "The Liabilities of Pension Funds"; J. F. Reilly: "Interpolation with Modified Coefficients"; H. L. Rietz: "Note on the 'Most Probable Number of Deaths' "; S. H. Pipe: "Modern History of Fraternal Insurance."

JOURNAL OF THE INSTITUTE OF ACTUARIES.

November, 1927. J. G. Baker and G. D. Stockman: "Casualty Insurance in the United States of America"; "An Illustration of Adjusted Differences"; "The Danish Writers on Interpolation."

GIORNALE DEGLI ECONOMISTI.

September, 1927. U. Ricci: "Ancora sull'offerta del risparmio"; A. Gulinelli: "L'Albania ed il suo riordinamento monetario"; R. Fubini: "Recenti pubblicazioni."

October, 1927. F. Savorgnan: "La composizione raziale della popolazione Americana"; G. Mortara: "Asservazioni sul commercio fra d'Italia e l'estero nei primi novi mesi del 1927"; L. Federici: "Gli investimenti industriali in regime di Svalutazione"; U. Ricci: "Discorrendo col Signor Bousquet."

November, 1927. G. Mortara: "Alcuni confronti internazionali in materia di rivalutazione monetaria"; F. A. Repaci: "Le vicende e le ripercussioni della legislazione finanziaria di guerra sull'ordinamento tributario italiano"; M. Pagano: "Il manifold system in relazione all'economia di personale e all'efficienza di lavoro nell'azienda bancaria."

GIORNALE DI MATEMATICA FINANZIARIA.

June–October, 1927. F. Insolera: "Dopo l'VIII Congresso degli Attuari"; G. Santacroce: "Su la rappresentazione analitica di una distribuzione di frequenze mediante una serie di Hermite-Charlier"; G. Usai: "Sul valore attuale di una rendita postecipata e temporanea"; E. Lenzi: "Sul rischio matematico di assicurazioni singole"; E. Del Vecchio: "Sulla teoria degli ammortamenti (cont.)"; L. Lordi: "Sulle annualità variabili"; C. Bussi: "Sul Montante sinusoidale."

JOURNAL DE LA SOCIÉTÉ DE STATISTIQUE DE PARIS.

October, 1927. M. G. Cadoux: "L'Avenir économique de nos colonies. Quelques statistiques"; M. L. March: "F. Y. Edgeworth et les indices des mouvements des prix"; M. J. Giard: "Nécrologie (Ernest Plocq)"; M. P. Ricard: "Chronique des banques et marchés monétaires."

November, 1927. M. M. Moine: "La répartition de la population mondiale recensée vers 1920, et l'influence de l'immigration enregistrée en France au recensement de 1921"; M. H. Bunle: "Chronique de démographie."

December, 1927. M. Ichok: "Le problème des Immigrés Malades."

MISCELLANEOUS REPRINTS AND PAMPHLETS

AN ANALYSIS OF COMPENSATED ACCIDENTS TO MINORS FOR THE YEAR 1924. Special Bulletin Number 17. 1926. Department of Labor and Industry. Harrisburg, Pennsylvania.

SPECIAL ASPECTS OF THE DECLINING TUBERCULOSIS DEATH-RATE IN THE UNITED STATES, by Louis I. Dublin and George H. Van Buren. Reprinted from *Tubercle*, October 1927.

SEASONALITY IN SILVER PRICES AND EXCHANGE RATES, by Dickson H. Leavens. Reprinted from the *Chinese Economic Journal*, September 1927.

INJURED WORKERS AND THEIR DEPENDENTS FINANCIALLY BENEFITED BY AMENDMENTS TO WORKMEN'S COMPENSATION ACT, by William J. Maguire. From *Labor and Industry*, May 1927. p. 14. Department of Labor and Industry, Harrisburg, Pennsylvania.

MORTALITY EXPERIENCE OF THE FIRST NINE MONTHS OF 1927. Metropolitan Life Insurance Company Statistical Bulletin, October 1927.

A HISTORICAL RETROSPECT ON EXPECTATION OF LIFE. Metropolitan Life Insurance Company Statistical Bulletin, November 1927.

DAS DEUTSCHTUM KONGRESSPOLENS, by Andreas Mückler. Schriften des Institutes für Statistik der Minderheitsvölker an der Universität Wien. 1927.

THE PROBLEM OF TRADE ASSOCIATION LAW, by David L. Podell and B. S. Kirsh. Reprinted from *St. Johns Law Review*, January 1928.

THE NATURE OF CYCLICAL FLUCTUATIONS IN ELECTRIC POWER PRODUCTION DATA. Bureau of Business Research, University of Illinois Bulletin 16.

OPERATING EXPENSES OF RETAIL GROCERY STORES IN NEBRASKA 1926. University of Nebraska Studies in Business No. 20.

STUDI DEL LABORATORIO DI STATISTICA, by Felice Vinci. R. Instituto Superiore di Scienze Economiche e Commerciali-Venezia.